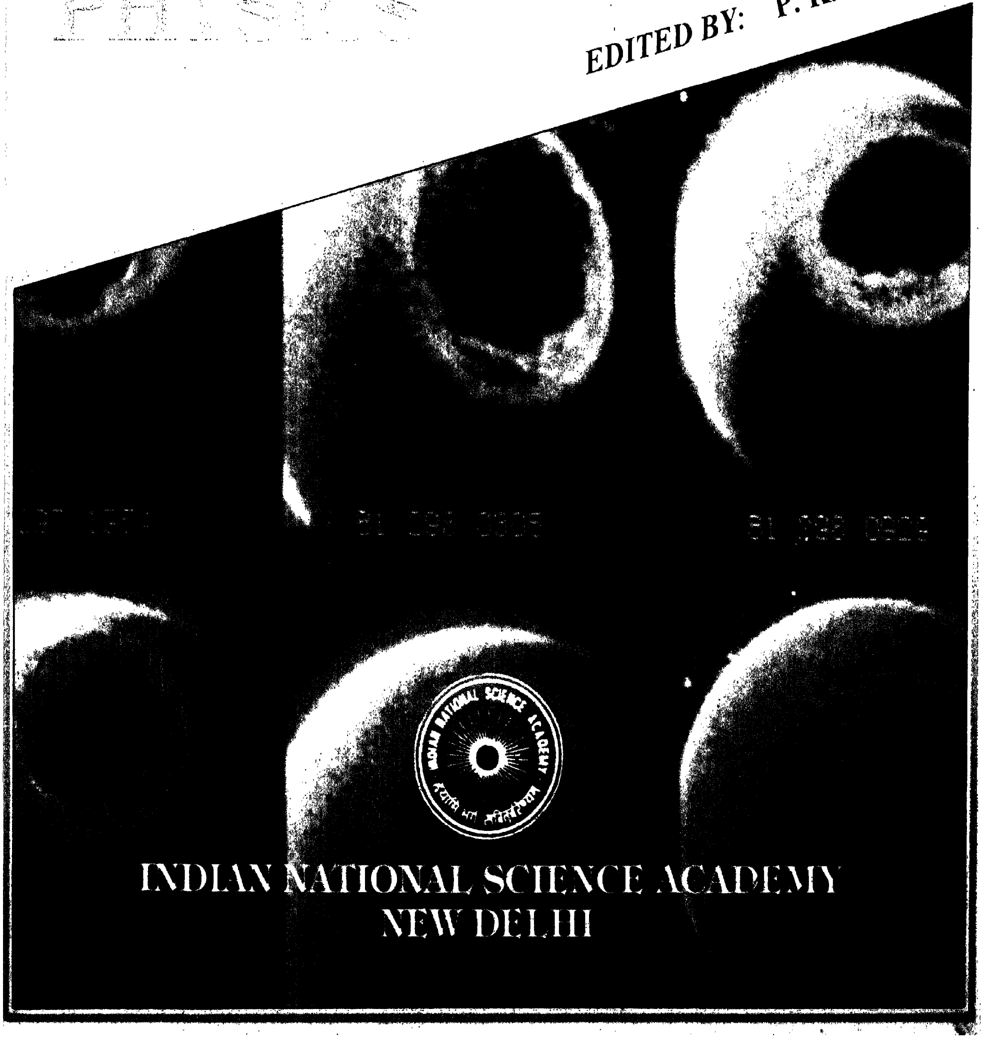
RECENT ADVANCES IN ATMOSPHERIC PHYSICS

EDITED BY: P. K. DAS



INDIAN NATIONAL SCIENCE ACADEMY
NEW DELHI

RECENT ADVANCES
IN
ATMOSPHERIC PHYSICS

Edited by

P. K. DAS

INDIAN NATIONAL SCIENCE ACADEMY
NEW DELHI
1994

© Indian National Science Academy, New Delhi

Editor of Publications: Professor O. P. Bhutani, FNA

Editorial Staff: Dr. M. Dhara, AES-I
J. Saketharaman, AES-II
M. Ranganathan, S.O.-III

First published in March 1994.

Price: Rs. 450/-; U.S. $ 150/-

Published by Shri S K Sahni, Executive Secretary, Indian National Science Academy, Bahadur Shah Zafar Marg, New Delhi 110002 and printed by Publications & Information Directorate, Dr K S Krishnan Marg, New Delhi-110012.

FOREWORD

The publication on "Recent Advances in Atmospheric Physics" coincided with the International Seminar on Numerical Weather Prediction in the Tropics. This Seminar was organised by the National Centre for Medium Range Weather Forecasting (NCMRWF) in New Delhi on January 31, 1994. Dr. S.K. Mishra the Director (NCMRWF) and Dr. U.C. Mohanty, the Jt. Director spared no pains to make the Seminar a success. Earlier, the contents of this publication were published in the form of a Special Number on Atmospheric Physics of the *Proceedings of INSA, Part A*. The warm response to this Special Number at the time of the International Seminar has encouraged us to bring out a separate publication, instead of just a Special Number, on the general topic of Atmospheric Physics. An additional paper on 'climate modelling' by L. Bengtsson and L. Dümenil of the Max Planck Institute in Germany has been added to this publication.

It has been our aim to provide the reader with a general view of the wide variety of problems in Atmospheric Physics. These problems, we feel, need further research to arrive at meaningful solutions. We hope readers will be encouraged to find other solutions by their own effort and research. This is in line with the basic tenet of INSA, that is, to encourage research and to develop expertise in different fields.

I wish to place on record our thanks to Dr. P.K. Das, formerly Secretary, INSA, for his initiative in compiling the present volume. Professor B.L.S. Prakasa Rao, former Editor of INSA, and Professor O.P. Bhutani the Current Editor, helped us to bring out the publication in time. We express our sincere thanks to all of them.

S.K. JOSHI
PRESIDENT, INSA

Dated: February 11, 1994

PREFACE

This *Recent Advances in Atmospheric Physics* is the outcome of a kind invitation to me from Professor B L S Prakasa Rao, FNA, the former Editor of Publications of the Academy's journal for Physical Sciences. It was our intention to bring out a volume that will concentrate on the current work in this important field within our country and abroad.

It was not very clear to me at the outset how to organise this work because 'Atmospheric Physics' is a wide and varied field. The important question in my mind was: Would it be preferable to focus on a few specialised areas or should we attempt a broad spectrum covering most areas where work is currently in progress in India and elsewhere. Eventually, the latter alternative seemed to be more preferable. Consequently, invitations were extended to a large number of scientists, both those living abroad and in India to contribute towards this Number.

The response was overwhelming; and I would like to place on record my gratitude to all who were generous enough to send in an article despite an imposition on their time. A special word of thanks is due to those scientists from abroad who were kind enough to send in a contribution, because their work will be a source of encouragement to our own younger scientists.

The contents have been divided into 5 broad categories, namely (i) Atmospheric structure and dynamics, (ii) Special systems, (iii) Natural calamities and numerical simulation experiments, (iv) The upper atmosphere and, lastly, (v) Remote sensing. I expect readers will find this Number interesting. Professor Newell, an acknowledged expert, has provided an interesting exposition of the temperature limits of the atmosphere. This is a well reasoned paper because it explains why one cannot assume an atmosphere to behave as a simple thermostat. Scientists from China have contributed papers to indicate the interesting work that is in progress in their country. And, there is a paper on Chaos and Predictability in Forecasting the Monsoon by T N Palmer. Of the Indian contributions, there are papers on numerical simulations of natural calamities, such as, tropical cyclones and a scheme by which real time prediction of storm surges becomes possible even on a personal computer. There is a similar paper on a critical study on long range prediction of monsoon rains with empirical orthogonal functions.

There are comprehensive papers by Dr Subbaraya and others on different facts of the Upper Atmosphere and Remote Sensing. In the latter, readers would be interested to find contributions on techniques that are being developed by the Indian satellite, namely, the INSAT series. We do hope that this would lead to increasing global cooperation in satellite technology, because this must be one of the most remarkable advances in technology of the present century. Exciting developments are more than likely in the years to come.

Finally, I would like to express my indebtedness to the staff of the Indian National Science Academy for the meticulous care they have taken in preparing this special number. I would like to thank Dr M Dhara, Shri J Saketharaman, Shri M Ranganathan and Mrs. Kiran Chawla for their very valuable help to bring out this special number, and for their cooperation and assistance which was given freely despite the limitations on their time. The point to be emphasised is that this was a team effort in which a large number of people, including the authors and the editorial and secretarial staff made valuable contributions. My special thanks are also due to Dr B M Reddy, an expert in Radio Physics and the Upper Atmosphere for his kind assistance. There were several Officers of the Publications and Information Directorate (PID) who helped to bring out this publication in time. In this context, mention is made of Shri V Ramachandran, Shri Kaushal Kishore, Shri Ganesh Sahani and Dr G P Phondke.

Date: February 11, 1994

P K Das, FNA
Guest Editor, INSA
A-59, Kailash Colony
New Delhi 110 048

CONTENTS

Page

Foreword *iii*

Guest Editorial *v*

I. Atmospheric Structure and Dynamics

Temperature Limits in the Climate System
Reginald E Newell and Zhongxiang Wu 1

A Study on the Dynamics and Energetics of the Indian Summer Monsoon **U C Mohanty and K J Ramesh** 23

Chaos and Predictability in Forecasting the Monsoons **T N Palmer** 57

Numerical Experiments of the Influences of Anomaly Heating over the Tibetan Plateau on the Formation and Evolution of the Asian Summer Monsoon **Fan Yun and Chen Longxun** 67

East Asian Summer Monsoon—A Case Study
Qing-Cun-Zeng, Bang-Lin Zhang, You-Lin Liang and Si-Xiong Zhao 81

A Brief Review of the Studies on the Summer Monsoon in China
Guo Xiaorong 97

Dynamical Predictability of Seasonal Monsoon Rainfall: Problems and Prospects **B N Goswami** 101

II. Special Systems

Evolution of the Boundary Layer over the Southern Ocean
M Chatterjee 121

Terrain Variance Spectra for Indian Western Ghats
K Srinivasan and N Ramanathan 133

III. Natural Calamities and Numerical Simulation Experiments

A Numerical Simulation Study of the Role of Sea Surface Temperature in Tropical Cyclone Evolution
D V Bhaskar Rao and S S V S Ramakrishna 139

A Real Time Storm Surge Prediction System: An Application to East Coast of India
S K Dube, A D Rao, P C Sinha and P Chittibabu 157

The Nor'Wester Scenario: The Role of Internal Gravity Wave
D K Sinha and G K Sen 171

Long Range Weather Forecasting Techniques in India Using EOF Approach **H N Srivastava and S S Singh** 181

Variational Data Assimilation with an Adiabatic Version of the LMD GCM using Adjoint Method **H C Upadhyaya, Sangeeta Agarwal and O Talagrand** 193

Dynamical Models for Quantitative Precipitation Forecasts in India **P N Sen** 205

Importance of Mountain Waves in Aviation and Weather Hazards Associated with It **U S De** 217

IV. The Upper Atmosphere

Mesosphere: Coupling Region Between the Upper Atmosphere and Middle Atmosphere **B H Subbaraya** 227

Tropical Ionosphere and Radio Communications **D R Lakshmi** 243

The Auroral Ionosphere—A Television Screen for Processes in the Distant Magnetosphere **Girija Rajaram** 263

V. Remote Sensing

Physics and Techniques of Remote Sensing for Weather and Climate Research **P C Pandey** 283

Remote Sensing and Vegetation **B L Deekshatulu and R K Gupta** 299

Satellites for Monitoring Climate Change: The Emerging Scenario **R R Kelkar** 335

Application of Satellite Data for Monsoon Research **A V R Krishna Rao** 349

Studies on Atmospheric Aerosols at a Coastal Station, Trivandrum **B V Krishna Murthy** 369

Remote Sensing of Cloud Precipitable Water Content using INSAT-1D Data **J K Mishra, O P Sharma and M D Joshi** 387

Role of Time Delay in Climate Research* **Kaustubha Bhattacharya** 401

Atmospheric Climate Modelling—A Review of Current Activities at MPI Hamburg* **L Bengtsson and L Dümenil** 407

*Belongs to Section I. Late receipt, hence added at the end of the volume.

TEMPERATURE LIMITS IN THE CLIMATE SYSTEM

REGINALD E NEWELL AND ZHONGXIANG WU
Department of Earth, Atmospheric and Planetary Sciences
Room 54-1824, *Massachusetts Institute of Technology, Cambridge, MA* 02139, *USA*

(*Received* 19 *August* 1993; *Accepted* 30 *August* 1993)

Physical processes in the climate system appear to limit maximum and minimum temperatures in various regions of the ocean and atmosphere. Possible limits in four regions are discussed here: the surface ocean in the tropical west and east Pacific, the tropical free troposphere, and the lower tropical stratosphere. The processes are deduced from recent data sets but show some consistency with data from the past ice age. The lower limit on surface temperatures at high latitudes during ice age conditions is also examined.

Key Words: Climate; Limiting Temperatures; Warm Pool; Ages; Volcanic Activity; Associated Temperature Change

Introduction

Maps of tropical Pacific surface temperature (Fig. 1) show a region in the tropical west Pacific where temperatures are close to 29°C throughout the year, following the sun at lag between the hemispheres. Why is this value 29°C and not 25°C or 35°C? Why does the tropical east Pacific also never exceed this value? Likewise the tropical east Pacific rarely goes below 18°C. What controls this limit?

During an El Niño the extra energy flux into the atmosphere from additional evaporation and subsequent condensation can, in principle, heat the tropical roposphere by several degrees but the actual temperature change observed is much smaller. What physical mechanisms act to limit the temperature change of the tropical troposphere?

The last three volcanoes large enough to significantly affect the atmospheric transmission of solar radiation have shown large differences in transmission and in mass of stratosphere aerosol yet the temperature changes produced in the stratosphere have been almost the same (5-6°C). What are the factors that act to limit the temperature change of the tropical lower stratosphere?

During an ice age, temperatures fluctuate over a fairly wide range as shown by oxygen isotope ratio in ice cores yet the lowest temperatures reached, corresponding to about -42% in $\delta^{18}O$ values in the Summit core in Greenland (GRIP, 1993), are usually the same. What are the mechanisms that act to limit these high latitude surface air temperatures?

These questions and associated ramifications are examined here.

Sea Surface Temperature Limits

The factors that control temperature change in the oceanic surface layer may be written:

Fig 1(*a*)

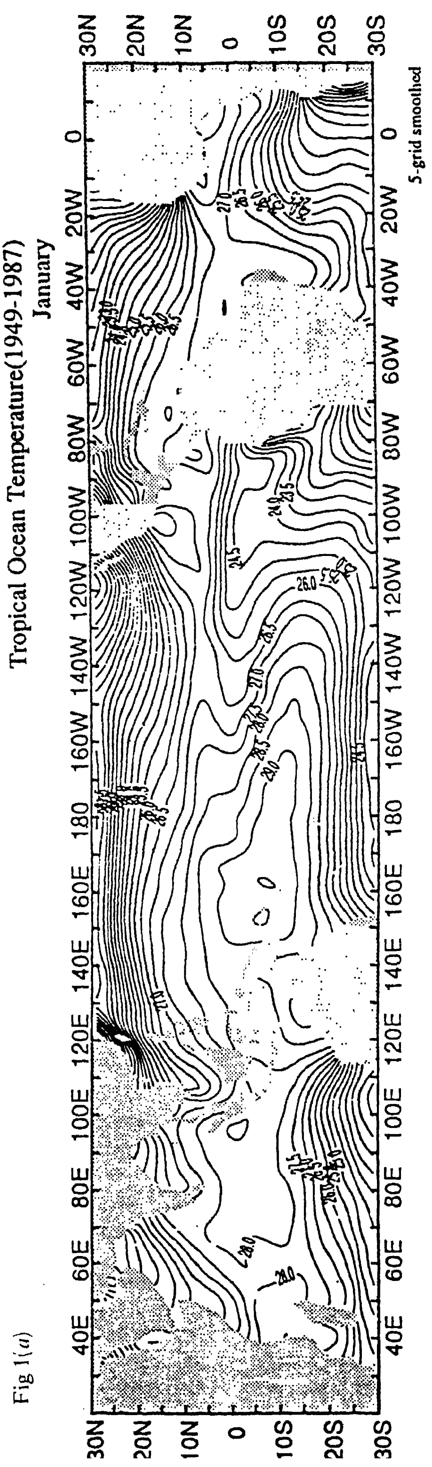


Fig 1(*b*)

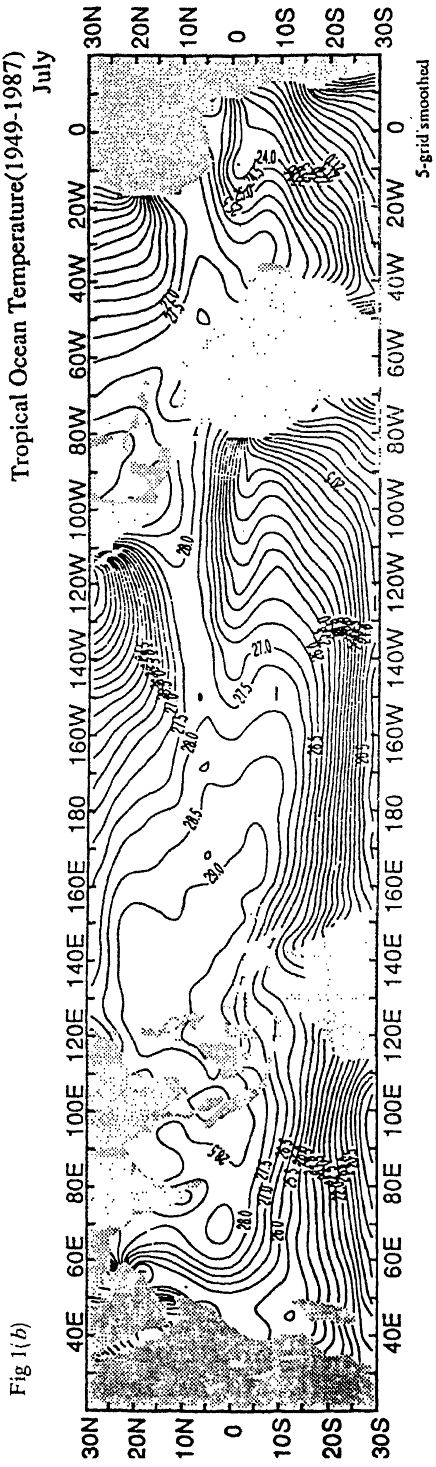


Fig 1 Mean values of tropical sea surface temperatures (°C), 1949-1987: *a* January; *b* July

$$\frac{\partial T_s}{\partial t} = \frac{Q_{sol} - Q_{IR} - Q_{SH} - Q_{LH}}{\int_{-D}^{0} \rho C dz} - u\frac{\partial T_s}{\partial x} - v\frac{\partial T_s}{\partial y} - w\frac{\partial T_s}{\partial z} - \frac{\partial}{\partial z} K_z \frac{\partial T_s}{\partial z}, \qquad \ldots (1)$$

where T_s is a surface layer temperature, u, v and w are oceanic flow components in the eastward, northward and upward directions (themselves denoted as x, y and z respectively) and K_z is a vertical eddy diffusion coefficient. The first four terms on the RHS denote the surface energy input by solar radiation Q_{sol}, the surface loss by infrared radiation Q_{IR}, the loss as sensible heat Q_{SH} and the loss as latent heat Q_{LH}. ρ represents the density of seawater and C the specific heat and D the thermal interaction depth over which the solar radiation is absorbed. T_s may be constant down to 200-300m in the west Pacific but only to 10m in the tropical East Pacific. One may in fact write the thermodynamic equation in terms of heat content changes and temperature of the isothermal layers but the same physics is involved. The fifth and sixth terms represent lateral advection while the seventh and eighth produce cooling by vertical advection (upwelling) or by mixing the surface layer with cooler water from regions below.

The surface energy components may be written as (Hsiung, 1986)[1]:

$$Q_{sol} = Q_0 Tr(1 - A)(1 - 0.62c + 0.0019\alpha), \qquad \ldots (2a)$$

$$Q_{IR} = \varepsilon\sigma T_a^4(0.39 - 0.05\sqrt{e})(1 - bc^2) + 4\varepsilon\sigma T_a^3(T_s - T_a), \qquad \ldots (2b)$$

$$Q_{SH} = \rho_a C_p C_H V(T_s - T_a), \qquad \ldots (2c)$$

$$Q_{LH} = \rho_a L C_e V(q_s - q_a) \qquad \ldots (2d)$$

and

$$Q_N = Q_{sol} - Q_{IR} - Q_{SH} - Q_L. \qquad \ldots (2e)$$

Symbols are defined as follows:-

Tr, atmospheric transmissivity; A, albedo taken from Payne[2]; b, the cloud coefficient taken from Budyko's atlas[3]; c, cloud amount in tenths; α, solar noon altitude in degrees; ε, emissivity of water taken as 0.97; σ, Stefan-Boltzmann's constant; e, water vapor pressure; T_a, temperature of the air; T_s, temperature of the sea surface; q_a, specific humidity of the air; q_s, saturation specific humidity at T_s; ρ_a, air density; C_p, specific heat of air; L, latent heat of evaporation; C_h and C_e, transfer coefficients for sensible and latent heat taken from Isemer and Hasse[4] and Large and Pond[5]; and V, wind speed.

Typical values for eq. 2 are shown in Table I for the Eastern and western tropical Pacific. Eq. 2*b*-2*d* are all functions of surface temperature. While surface temperature does not appear explicitly in equation 2*d*, q_s is a function of T_s being proportional to e_s/p, where p is surface pressure and e_s the saturation vapor pressure which is a function of temperature through the Clausius Clapeyron relationship. Typical conditions in the western Pacific warm pool are used to illustrate the dependence of the terms upon temperature in Fig. 2. The situation is simplified as the contributions by the second term of equation 2*b* are ignored. They are always very small (< 10 Wm^{-2}) with the typical values of $T_s - T_a$ that are observed (~ 1.0°C) (see Bottomley *et al.*, 1990)[6]. The mean

Table I

Components of the heat fluxes in different phases of El Niño events

	TWP (5S – 5N, 150-160E)		TEP (5S – O, 100 – 90W)	
	El Nino	La Nina	El Nino	La Nina
Latent Heat (W/M**2)	115	113	77	36
Sensible Heat	6	5	2	0
Longwave Radiation	44	47	47	48
Solar Radiation	195	206	200	204
Net Heat Fluxes (Ocean)	33	42	74	121
Mean SST (C)	29.1	29.3	25.1	22.1
Mean SST (C)	29.2		23.5	
Standard Dev. (C)	0.5		2.3	

Periods of each phase (year month)

El Nino:

1957.4-58.1 65.5-66.1 72.6-73.2 76.6-76.12 82.8-83.8 86.11-87.6

La Nina:

1964.4-64.7 67.9-68.1 70.6-71.2 73.11-74.3 75.10-76.1 78.3-78.10
1935.1-85.9

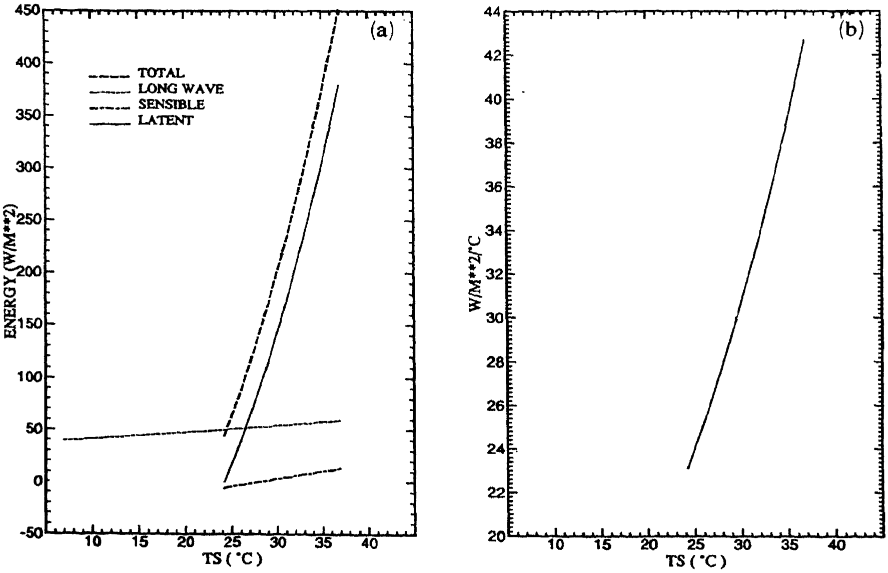


Fig 2 (*a*) Surface energy flux components for representative warm pool values ($v = 5$ m/sec; air temperature = 28°C; relative humidity = 80%); (*b*) sensitivity to temperature change wind speed is taken as 5 $msec^{-1}$, air temperature as 28°C and the atmospheric relative humidity is assumed to be 80%. For this region the temperature is often constant to 200-300m, as noted, so $\partial T/\partial z \rightarrow 0$ and, as is evident from the maps of Fig. 1, horizontal gradients are also very small. The first four terms of eq. (1) are then in approximate balance as observed. Fig. 2 shows that the equilibrium temperature is close to 30°C and that the sensitivity, obtained from the slope of the curve of energy versus temperature and shown in the right hand side of Fig. 2, is very low—typically it requires 30 Wm^{-2} to change the temperature by 1°C. This is thought to be due to the evaporational buffering introduced by Q_{LH} which is the largest term and therefore dominates this slope[7]. Fig. 2 is an updated version of similar graphs presented earlier[8,9]. It suggests that if additional radiative energy reaches the surface, for example by a decrease in cloudiness or an increase in atmospheric CO_2, it is offset in the warm pool region by additional evaporation. Note that the sensitivity of surface temperature to changes in incoming radiation would be much greater if infrared radiation provided the main balance during changes; the slope of the Q_{IR} curve is about 1 $Wm^{-2}°C^{-1}$.

Time series of sea surface temperature (SST) and the terms in equations 2 derived from the Comprehensive Ocean Atmosphere Data Set (COADS) are illustrated in Fig. 3 for the tropical east and west Pacific. The wind speeds used are included in Fig. 3*g*. It is clear that the sea temperature is quite stable in the tropical west Pacific showing little response to seasonal fluctuations in radiative input. The energy balance is between solar radiation and evaporational and infrared losses. There is reason to believe that the evaporational loss may be underestimated because when ships report calm winds there are often slight local flows accompanying convection that can encourage evaporation; approximately 25 Wm^{-2} may be added to Q_{LH} to take this process into account. As seen from Table 1 this reduces the net energy flux at the surface to close to zero (cf. Newell, 1986)[10]. A detailed consideration of this extra loss at low wind speeds has recently been given by Godfrey and Beljaars[11]. Conditions during El Niño are of considerable interest and it is clear from Fig. 3 that the waters of the tropical east Pacific receive less energy from the sun (because of more clouds—Fig. 3*b*) and lose more energy by evaporation during the 1982-83 El Niño than in the period immediately before and after (Fig. 3*d*). The water there warms towards the evaporational limit temperature (see Fig. 3*a*) of the tropical west Pacific[10]. In this region the upwelling terms are normally quite important and can account in general terms for the marked seasonal cycle[12]. But when upwelling and the other oceanic advection terms diminish as happens during El Niño the surface temperature in the east Pacific becomes more influenced by the surface flux terms[10] and rises towards the same limit that prevails in the west. Notice from Fig. 3*f* that the total surface flux tends towards zero in the 1982-1983 period, a condition normally more representative of the west Pacific.

The ocean surface temperature in the tropical west Pacific thus seems to be controlled by evaporational buffering with the same situation governing the maximum temperature achieved in the east Pacific during El Niño situations. By subtracting monthly mean values from 30°C, a maximum anomaly map for

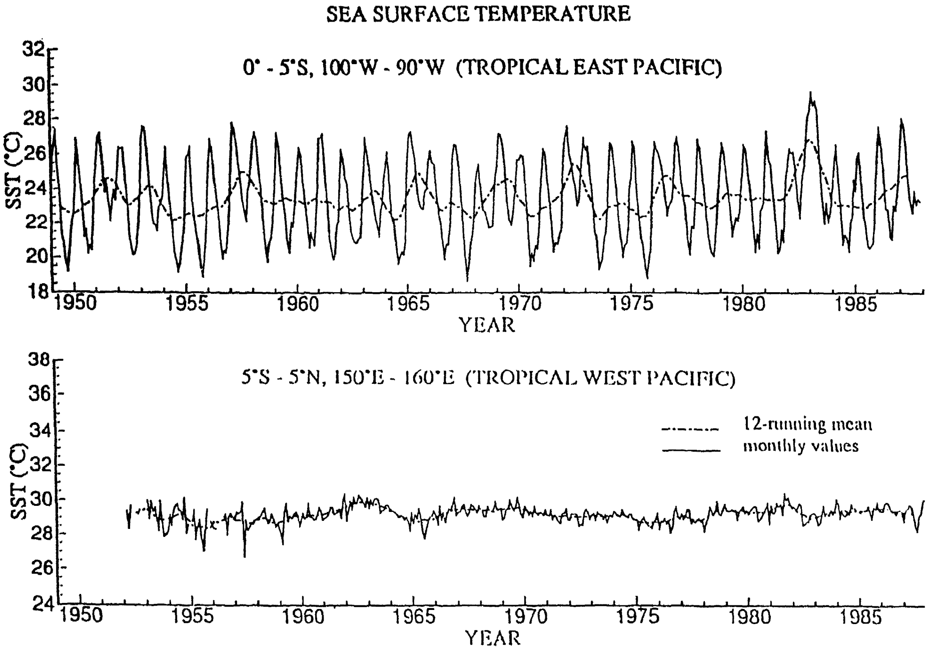


Fig 3(*a*)

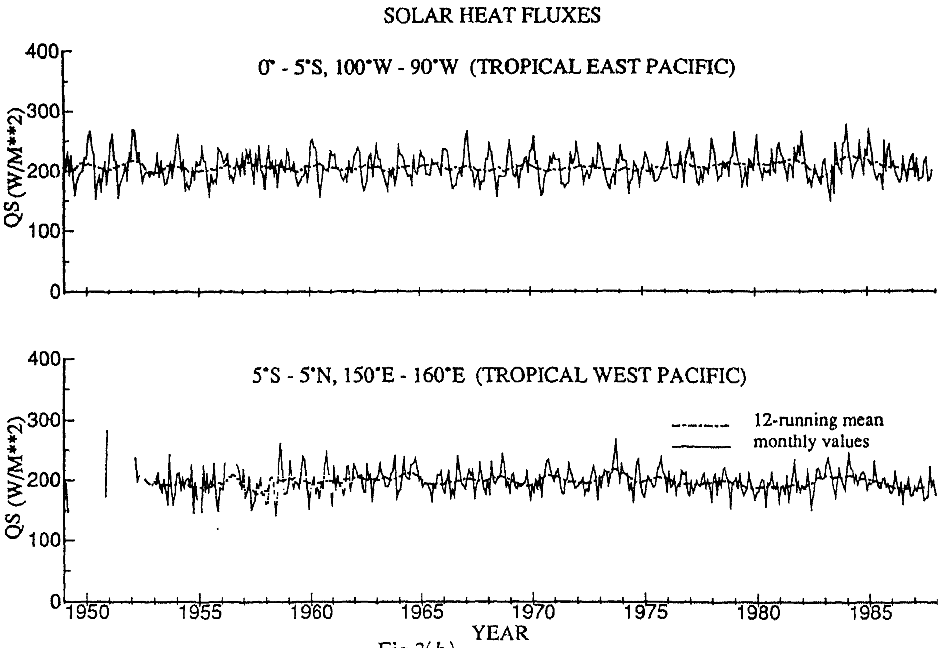


Fig 3(*b*)

Fig 3 Sea surface temperature, wind velocity and components of surface energy flux for the tropical east and west Pacific (*a*) sea surface temperature (*b*) incident solar radiation (*c*) sensible heat (*d*) latent heat (*e*) long wave radiation (*f*) net heat flux (*g*) wind speed. Long ticks are January [Figs 3(*c*), (*d*), (*e*), (*f*) and (*g*) are seen in the following pages]

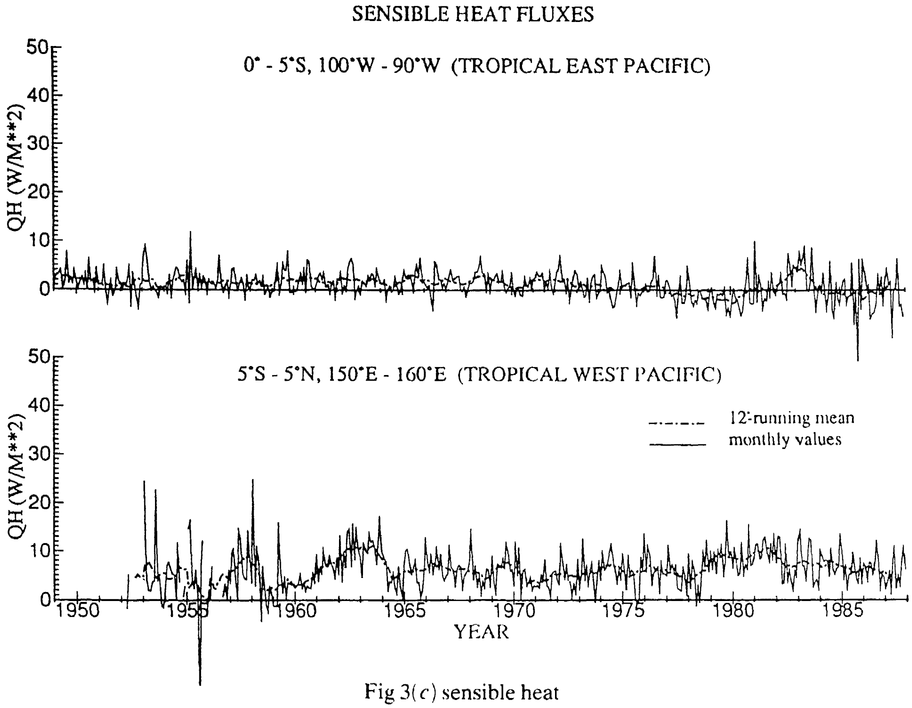


Fig 3(*c*) sensible heat

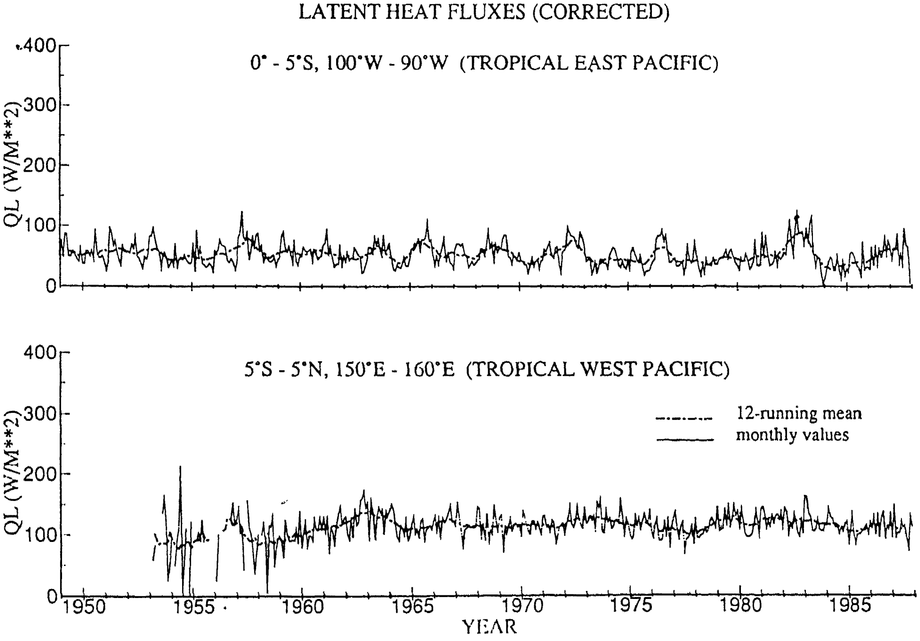


Fig 3(*d*) latent heat

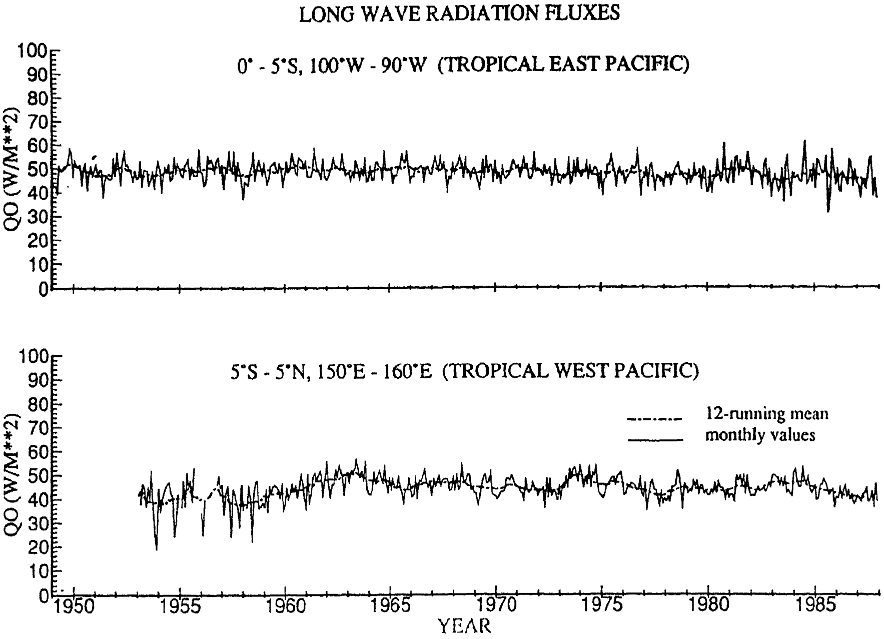


Fig 3(*e*) long wave radiation

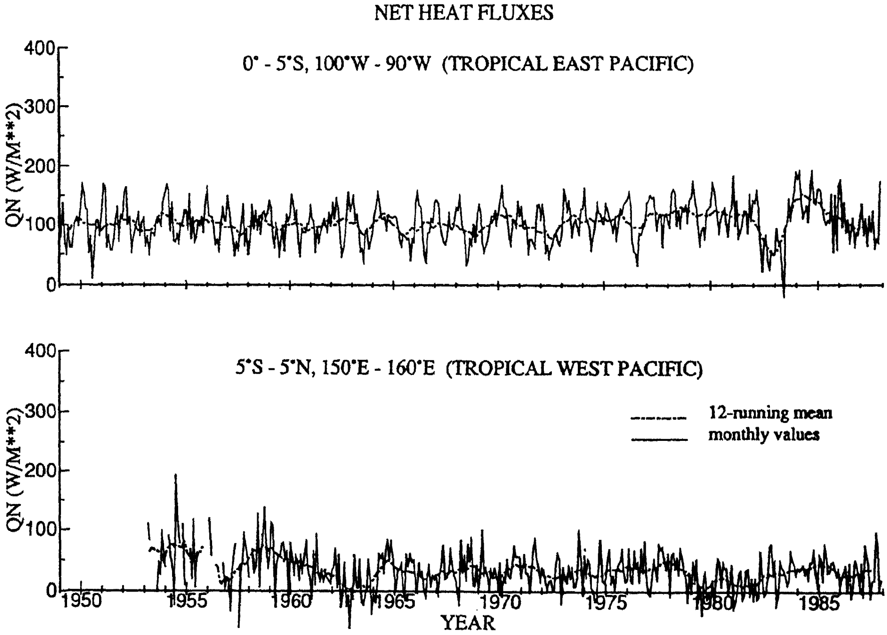


Fig 3(*f*) net heat flux

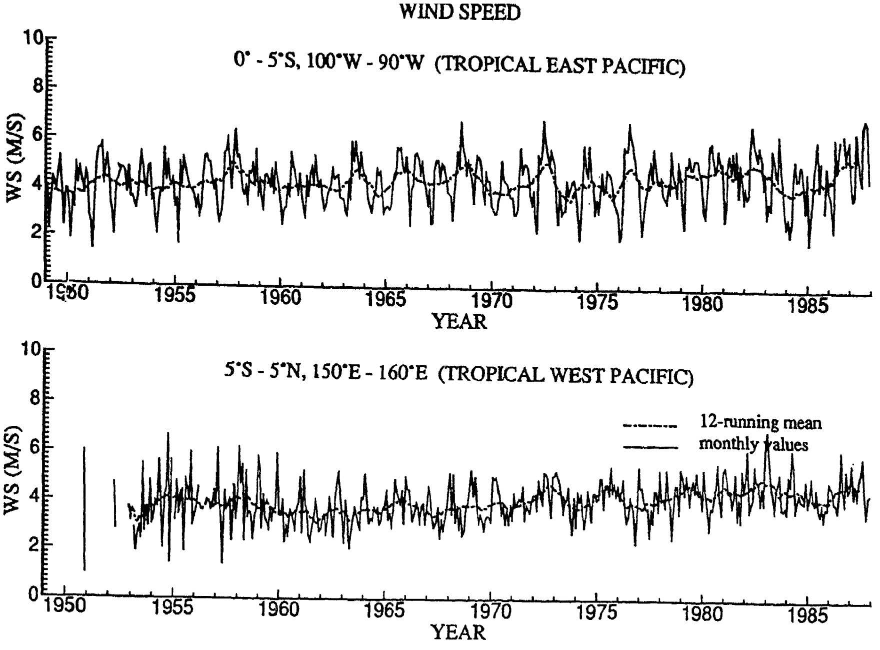


Fig 3(g) wind speed. Long ticks are January

each month can be reconstructed[13]; the actual anomalies that occurred in 1982-83 (up to ~6°C) are quite close to these maximum values. It is noteworthy that it is the Clausius Clapeyron relationship which controls the maximum temperature of the tropical ocean and therefore of the air in contact with the ocean, rather than the Stefan-Boltzmann law which is often invoked in simplified models. The evaporational buffering will also limit the minimum temperatures achieved in the warm pool region. Decreased solar radiation at the surface caused by greater than average cloudiness would be accompanied by decreased latent heat loss.

The limit on the minimum temperature in the tropical east Pacific is controlled by the balance between the net surface energy flux and the oceanic advective terms. Except during a strong El Niño the surface energy balance is about 100 Wm^{-2} (*see* Table I). The net contribution of the advective terms is quite complex; each is likely to be important. An excellent discussion of the energy budget of this eastern tropical Pacific region has been provided by Wyrtki[14]. Under the influence of a wind stress towards the west there are Ekman drifts away from the equator. These are accompanied by upwelling which brings up colder water from below which is then spread polewards by the Ekman drifts. In addition, the westerly stress builds up sea level towards the western Pacific and the consequent zonal sea level gradient has associated with it geostrophic drifts towards the equator thereby contributing to warm water advection. Zonal currents will also provide cold water advection in the east. We suggest here that as the wind increases there may be a minimum water temperature set by a balance between these advective terms. Stronger winds would increase upwelling and therefore cooling but at the same time increase the zonal sea level difference and thereby the equatorward geostrophic drift and the associated advective warming. It remains to work this hypothesis out with data; we have only examined the surface flux terms and the upwelling terms to show that their combination gives a reasonable explanation for the seasonal cycle[12].

A summary of the five terms in Fig. 3 divided into El Niño and La Niña phases is shown in Table I for the two regions. Incorporation of the factor to modify low wind speed situations as noted above adds about 25 Wm^{-2} to the latent heat loss and reduces the net flux in the west Pacific to close to zero, a condition invoked for Fig. 2. This condition would also be expected to apply to the east Pacific when SST there moves towards its equilibrium value during El Niño[10], an example of which appears in Fig. 3*f* in 1983.

Comparison of SST Evaporative Upper Limit with Other Ideas

Graham and Barnett[15] have also investigated warm pool SST limits and show that deep convection sets in above 27.5°C and that further SST increases have little effect on the intensity of the convection. Their work shows that large-scale divergence controls the deep convection at the higher temperatures. At the highest SST values convection is reduced according to their work, and this fits with our finding that in the Indian Ocean peak SST values (see histograms in Newell *et al.*, 1978)[7] occur in May when there is large scale atmospheric subsidence and divergence before the monsoon starts.

A different view was provided by Ramanathan and Collins[16] who suggest that highly reflective cirrus clouds shield the ocean from solar radiation and limit SST to less than 32°C. The problem with this explanation is that the warmest values appear where there is no cloud (see discussion below). If cloud is the limiting factor the SST should attain even higher values without cloud.

Wallace[17] argues that it is deep convection and not evaporation that provides the thermostat on sea temperature. He suggests that deep convection will rapidly carry away any surplus energy from "hot patches" and believes the skewness in the histograms of SST, with a sharp cut off at higher temperatures, is due to this effect. Ramanathan and Collins[18] have recently criticized Wallace's paper but seem to be mostly directing their remarks at evaporation as an explanation and not deep convection. They claim that evaporation decreases with increasing SST and that the evaporative heat flux decreases towards the warm pool. Both of these points are contrary to the computations of Table I or Fig. 3*d*. Waliser and Graham[19] have developed the work by Graham and Barnett[15] and they carry through the arguments in more detail, illustrating from data that the maximum temperatures occur under relatively cloud-free skies while the maximum of the SST population near 28°C is due to the reduction in the surface solar radiation by clouds associated with organized convection. They suggest that the thermostat is essentially the stability of the atmospheric column relative to large scale moist convection and solar flux. They imply that the cloud free limit is set by evaporation. Fu *et al.*[20] have recently joined the debate and from cirrus cloud observations suggest that cirrus cloud does not provide the main limit; they favour evaporation.

Tropospheric Free Air Temperature Limits

Studies of correlation patterns between global non-seasonal air temperature anomalies measured from space with microwave sounding units (MSU) and sea surface temperature anomalies show that over the last decade tropical air temperature changes throughout the entire tropical belt between about 25°N and 25°S are related to sea temperature changes in the tropical east Pacific; there is no relationship with the much smaller temperature changes in the tropical west Pacific[21]. The vertical extent of the atmosphere covered by the microwave sounder is roughly from the surface to 7km[22]. The lack of relationship between the two media in the west is thought to be due to the observation that the west Pacific is near to its evaporational limit; even though the actual evaporative loss is large, as is evident from Fig. 3*d*, the changes, unlike those in the east Pacific, are not systematically related to SST. The tropical air temperature changes by about 0.7°C per °C change of sea temperature in the tropical east Pacific according to these recent studies; this compares with previously reported values of about 0.5°C[13,23,24]. The tropical east Pacific is also thought to control global atmospheric concentration[7] of atmospheric CO_2.

The equation governing air temperature change may be written

$$\frac{\partial T_a}{\partial t}=\frac{Q_{RAD}+Q_{LH}+Q_{SH}}{\rho_a C_p}-\Gamma\omega-\frac{v\partial T}{\partial y}-\frac{u\partial T}{\partial x}, \qquad \ldots(3)$$

where eddy flux convergence terms, which are small in the tropics, are ignored. u, v, and ω now refer to atmospheric motion with $\omega = dp/dt$ and ρ_a is the density of the air. Γ is a stability factor $\frac{RT}{pC_p}-\frac{\partial T}{\partial p}$ where p is pressure and R is the gas constant.

The first question is—which of these terms is of most importance in producing the tropical belt of increased temperature that accompanies El Niño? From Fig. 3 we note that additional evaporation injects water vapor into the tropical belt. Using a model of the tropical atmosphere developed by Gill[25] the second author has modeled the tropical atmosphere for the case where additional latent heat of about 40 Wm^{-2} is injected into the eastern Pacific and finds that an increase of temperature of about the observed size (~ 1.0°C) occurs by adiabatic subsidence (the $\Gamma\omega$ term). Another possibility is that there is a change in the meridional heat transport from the tropics to middle latitudes; the appropriate formulation may be obtained from the last two terms of eq. (3) but the data necessary to study this possibility is not presently available here. The model produces a warm tropical strip but it seems that some other factor is involved to offset the adiabatic subsidence and limit the troposphere temperature increase.

The term Q_{RAD} incorporates at least five sub-components: heating by visible and near infrared (0.6μm-4μm) radiation absorbed by gases and aerosol and cooling by infrared emissions from water vapor, carbon dioxide and ozone. The tropospheric variable which changes most in the near infrared portion of the spectrum is water vapor. The additional water vapor given off by the tropical eastern Pacific during El Nino causes additional heating through enhanced near infrared absorption as well as condensation. But the heating is offset by additional infrared cooling which is also dominated by water vapor. Over most of the troposphere above the boundary layer increased water vapor is associated with increased infrared cooling rates[26]. The horizontal temperature gradients are relatively small in the tropics. Thus it seems likely that these radiative mechanisms, coupled with the latent heat liberation, may together produce another limit to atmospheric temperature rise. To study the role of changing water vapor in the east Pacific on the near infrared heating rates and the infrared cooling rates we have used a radiative heating rate program, designed by T Dopplick and described and revised by Hoffman[27], applied to wa-

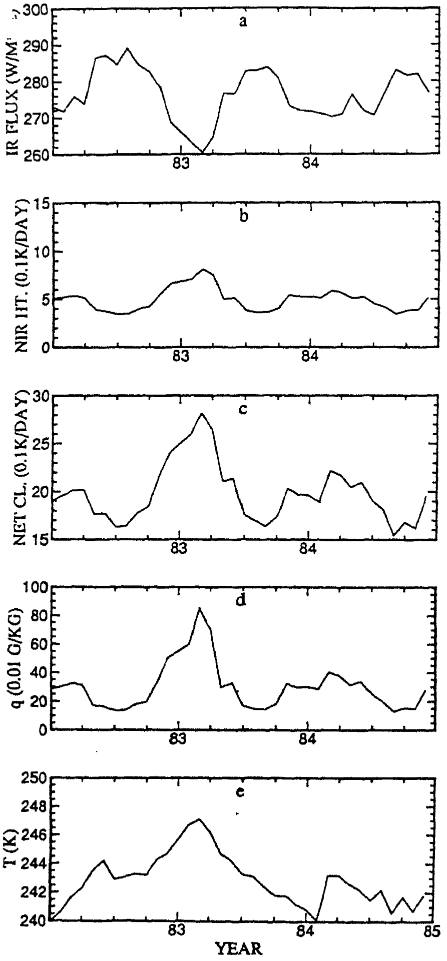


Fig 4 Heating rates for Atuona (9°40'S, 139°02 W) based on mean monthly radiosonde data: (*a*) infrared flux at 200 hPa; (*b*) heating by near infrared absorption at 300 hPa; (*c*) net cooling by infrared radiation at 300 hPa; (*d*) specific humidity at 300 hPa; (*e*) observed temperature variation at 300 hPa, 1982-1984. Long ticks are January

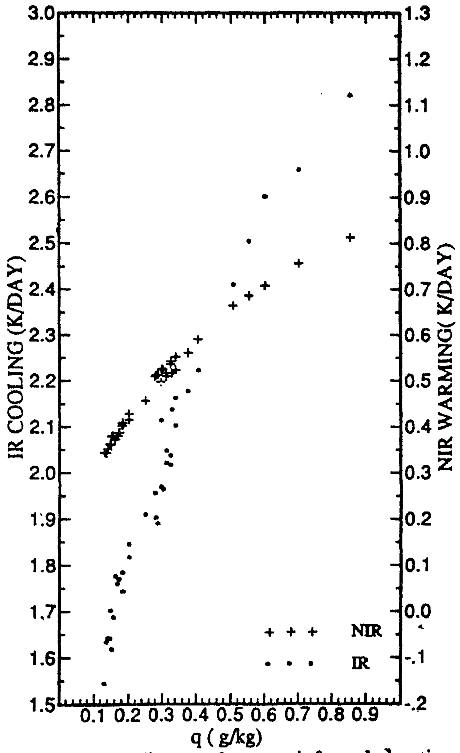


Fig 5 Infrared cooling and near infrared heating rates with water vapor mixing ratio at 300 hPa, based on 1982-1986 period

ter vapor and temperature values taken from monthly mean soundings for Atuona (140°W, 10°S) for the 1982-84 period. These calculations, performed for another purpose, assumed cloud-free skies (M. Lozano, *private commun.*). Values of temperature and specific humidity at 300 hPa are shown in Fig. 4 together with the infrared cooling, the near infrared heating and the infrared upward flux at 200 hPa. As already noted when the eastern tropical Pacific warmed in the 1982-83 El Niño the specific humidity increased and with it the near infrared heating and the infrared cooling. As expected the upward infrared flux also decreased. Obviously there would be concomitant changes in the horizontal energy fluxes to be taken into account if we were trying to balance the energy budget for the tropical belt; here we concentrate on the local cooling rate.

The near infrared heating rates and total infrared cooling rates for 300 hPa are shown in Fig. 5 as a function of specific humidity. At a value of about 0.5 g kg^{-1} the sensitivities of these rates to changing water vapor are 0.6°C $\mathrm{day}^{-1}/\mathrm{g\ kg}^{-1}$ and -1.8°C $\mathrm{day}^{-1}/\mathrm{g\ kg}^{-1}$ respectively. The infrared cooling dominates and provides a buffering process for the temperature changes. The additional evaporation and rainfall is not known very well and needs careful study before this temperature limit can be established. It is important to recognize that during El Niño, when additional evaporation from the eastern

Tropical Pacific occurs, the condensate can be carried by the zonal winds throughout the tropical belt and also to higher latitudes even though the largest water vapor anomalies occur in the eastern Pacific. The heating rate from the additional condensation can be quite large (~ 0.5°C day^{-1}) but the full value is not realized because of inrared cooling. In general, the heating by condensation occurs over a region with a smaller latitude span that the region in which heating by subsidence occurs. In order for a net temperature increase to occur the over subsidence heating must outweigh the additional radiative cooling or the meridional energy flux to higher latitudes must decrease.

In the seasonal cycle water vapor has a semi-annual cycle in the equatorial zone[28] and the tropical microwave temperatures also show a semi-annual temperature maximum in the same phase (W. Hu, *private commun.*) which may be produced by near infrared absorption. The additional heating during El Nino will thus be composed of condensation heating from rainfall, near infrared solar absorption, and warming by subsidence. Numerical values of a possible limit remain to be worked out from a more comprehensive set of grid point data which includes clouds.

Lower Stratospheric Temperature Limits

Another region where a natural temperature limit may exist is in the tropical lower stratosphere. This region is sometimes invaded by volcanic aerosol and

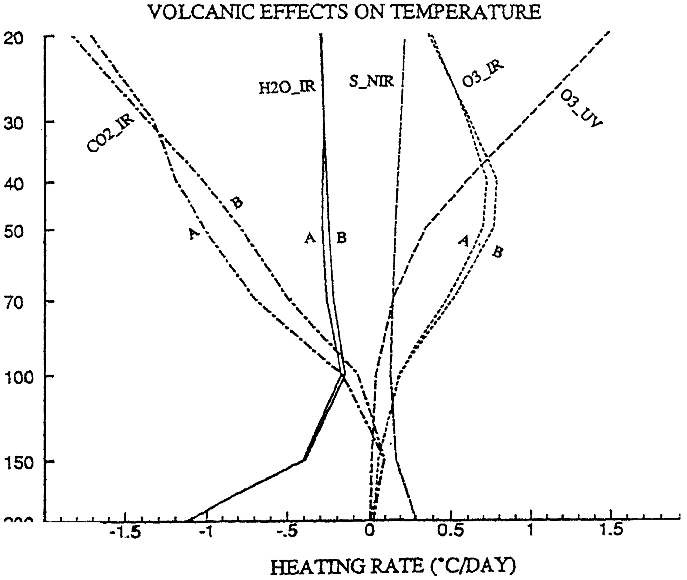


Fig 6 Radiative heating and cooling rates for tropical radiosonde sounding for Guam for conditions before (*B*) and after (*A*) the eruption of Mt. Pinatubo in June, 1991. S NIR is near infrared heating by water vapor, molecular oxygen and carbon dioxide; H_2O IR and CO_2 IR represent infrared cooling by these gases, O_3 UV is heating by ozone absorption of solar radiation, and O_3 IR represents infrared heating. At 50hPa, A temperatures are higher than B by 6°C

absorption by this aerosol of sunlight and perhaps, in some cases, of upwelling infrared radiation causes temperature increases. According to radiosonde data, these measured about 6°C after the Mt. Agung eruption[29] in March 1963 and about 5°C after El Chichón[30,31] in April 1982. Recent estimates based on MSU data (Christy, *private commun.*) show that the Pinatubo and El Chichón eruptions gave similar tropical stratosphere increases. The amounts of dust in the stratosphere from those eruptions has differed widely with Pinatubo thought to produce the largest amounts. It is therefore suggested that there may be another temperature limit in this region. Using the same radiation formulation as that discussed above, it can be seen from Fig. 6 that atmospheric CO_2 in the tropical lower stratosphere produces cooling rates at the higher post-volcanic temperatures (the difference was assumed to be 6°C at 50 hPa) which tend to offset heating by aerosol absorption[32]. The infrared cooling from CO_2 increases with temperature and in this case it is essentially the Stefan-Boltzmann law that is controlling the limit. It is the component of the first term of equation (3), Q_{RAD}, that provides the balance as seen in Fig. 6, although changes in vertical motion may also be involved.

Possible Limits on Ice Age Temperatures

How do these limits fit with what is known about conditions during the past ice age? Temperatures in the tropical eastern Pacific were lower than the present by 4-6°C at 18000 B.P. in the original estimates[33] but only about 2°C in more refined estimates[34]. These values need more work before they can be compared with present values of seasonal SST changes such as those in Fig. 3*a*. Pollen data from Africa[35] and South America[36] show descent of the tree line of 1000-1100m and 1200-1500m respectively in these two regions, corresponding to temperature decreases of more than 6°C. Because temperatures were generally lower the atmosphere was able to hold less water vapor. Specific humidity values for the past ice age were scaled from present values with the assumption that the early CLIMAP surface temperature results were valid and that relative humidity remained the same; temperature lapse rates were also taken to be the same as present values. Radiative cooling rates were computed for these synthetic ice age conditions and the assumption made that the first three terms on the right hand side of equation 3 summed to zero[37]. It turned out that the cooling rates decreased by about 10% relative to present values and the conclusion was made that the latent heat term and therefore rainfall must also have decreased by this amount. This point was in reasonable agreement with ice age data (e.g. Flint, 1971)[38] and with findings from general circulation models (e.g., Gates 1976 found a decrease of 14%)[39].

We have argued already that the tropical west Pacific warm pool temperature is essentially stable and that the tropical east Pacific surface temperature is limited on the lower side. Similar arguments may be made for the warm regions in the Indian Ocean and the upwelling in the Atlantic. The implication is that events in middle or high latitudes must be responsible for the lower limit on ice age temperatures. Latent heat liberation is the main factor in eq. (3) providing energy to the atmosphere. As noted above, when the atmosphere cools this term diminishes. But as high latitude cooling increases meridional temperature gradients steepen and overall wind velocities increase[40]. At present largest evaporation occurs off the east coast of the continents in winter as cold dry air streams out over the ocean. With stronger winds and ice present on the continents throughout the year in the ice ages, there will most likely be a minimum value for this evaporation term and therefore a lower limit on Q_{LH} and temperature in eq. (3).

This possible lower limit may be treated from another perspective by considering the meridional fluxes of energy by the atmosphere and ocean. When one or both of these fluxes diminishes below its long term value there is a resulting energy deficit at high latitudes and ice may build up; conversely the ice would diminish when the fluxes are above present values. The change in the amount of ice in the polar cap regions can be used to assess the associated change in the fluxes[41]. The time assumed for ice age build-up in our previous work was about 8000 years. A more modern view may yield a higher value by about a factor of two and this gives a value for the required flux changes of about 2.5 × 10^{14}W. Present day fluxes are shown in Fig. 7. At 30°N atmospheric fluxes (taken here from Oort and Vonder Haar, 1976)[42] are a maximum in the northern hemisphere winter while oceanic fluxes (taken here from Hsiung *et al.*, 1989)[43] are maximum in northern hemisphere summer. The radiation budget which these fluxes balance and the relative roles of heat stored in the ocean and meridional energy fluxes in maintaining winter temperatures is discussed elsewhere[8]. Our previous suggestion was that a decrease in the oceanic energy flux initiates a change towards an ice age[41]. A little more is now known

Fig 7(*a*)

(*a*)

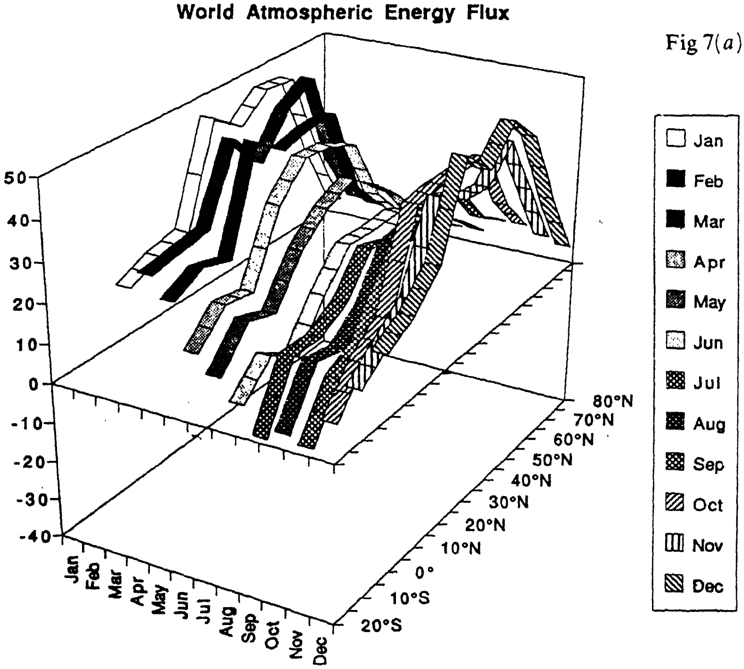
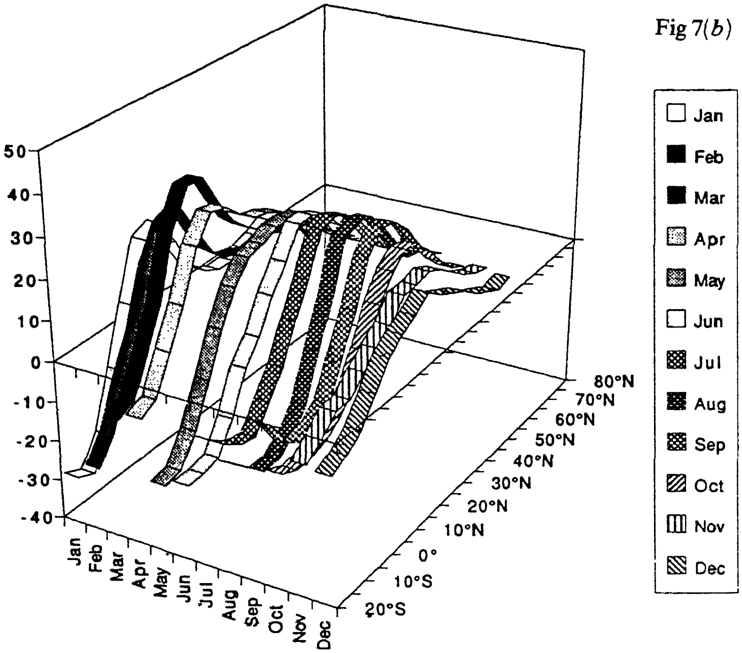


Fig 7 Meridional world energy flux (units 10^{14} W): (*a*) atmosphere;·

World Oceanic Energy Flux

(*b*)



Sum of World Oceanic and Atmospheric Energy Fluxes

(*c*)

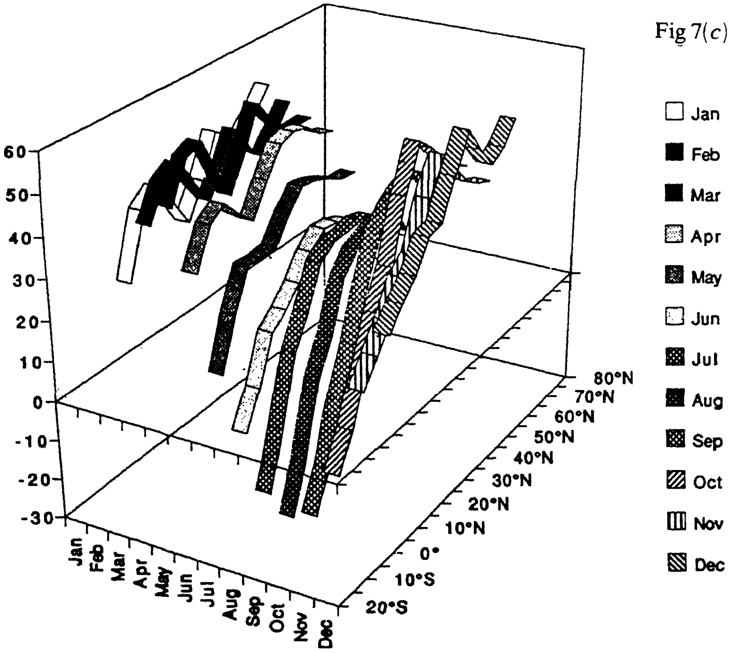


Fig 7 Meridional world energy flux (units 10^{14} W):
(*b*) ocean; (*c*) sum of atmosphere and ocean

about the annual cycle of the oceanic fluxes for each ocean separately (Hsiung *et al., loc. cit.*)[43] and about the possible linkage of the Indian and South Atlantic Ocean (Gordon, 1986)[44]. It is thought that the Indian Ocean supplies energy to the South Atlantic which then crosses the equator into the North Atlantic. Thus the place to look for changes in solar radiation that may influence ice formation is not necessarily high latitudes in summer but may be tropical latitudes whenever there is an energy transport from the Indian Ocean into the Atlantic. The breakdown of energy flux components appears in Figs 8. Clearly the Indian Ocean carries energy southwards and the Atlantic northwards in April, July and October but the reverse is true in January. This linkage may be the best place to look for the influence of Milankovitch solar radiation variations on climate. If the oceanic energy supply to the North Atlantic is the ultimate mechanism that initiates ice ages, the fact that the atmospheric energy flux increases with the meridional temperature gradient (*see* Newell, 1974 and Fig. 7*a*)[41] suggests that at some point the atmospheric flux compensates for the

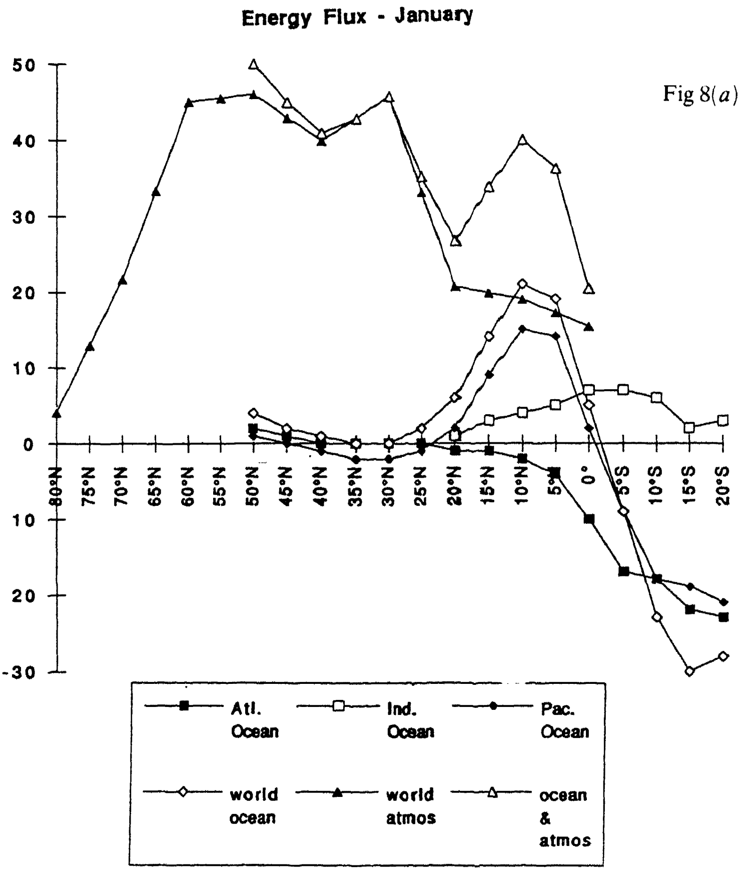


Fig 8 Individual components as indicated of meridional energy flux (units 10^{14} W): (*a*) January;

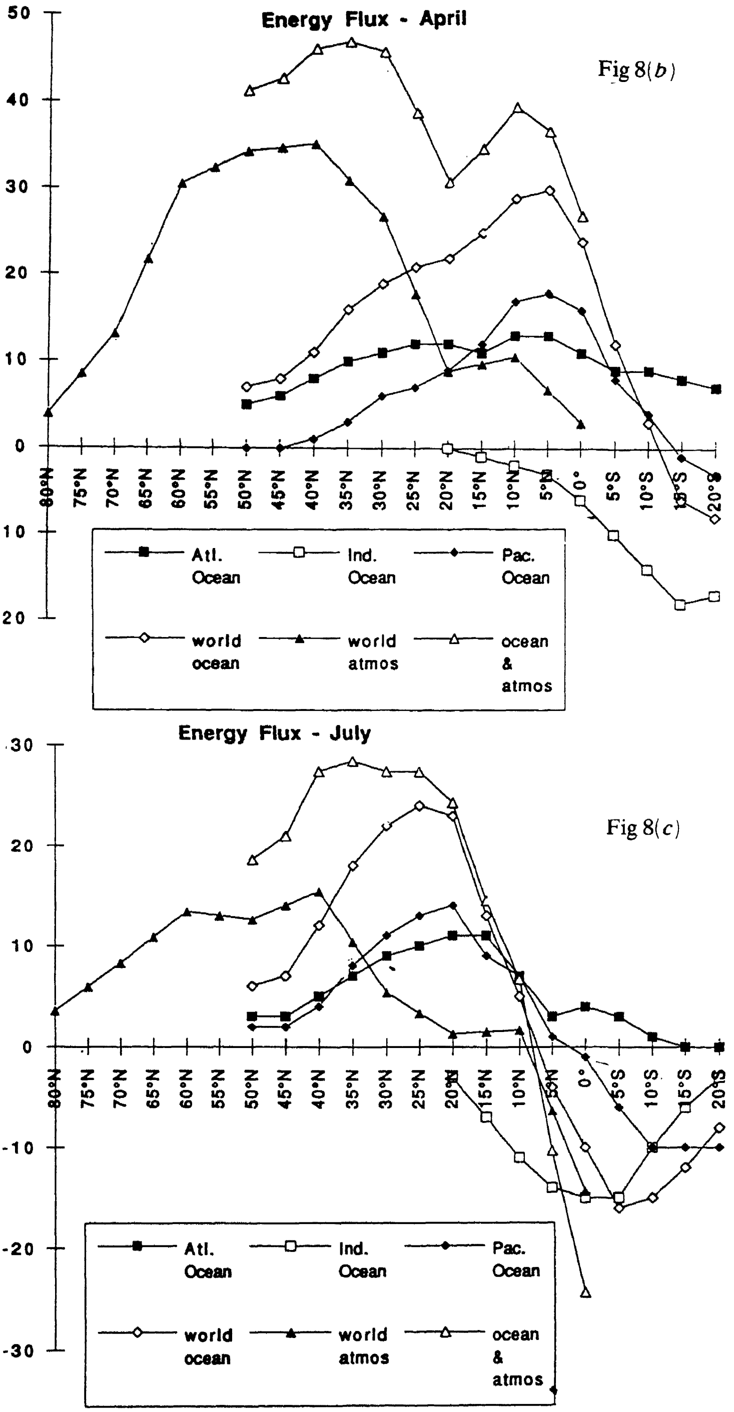


Fig 8 Individual components as indicated of meridional energy flux (units 10^{14} W): (*b*) April; (*c*) July

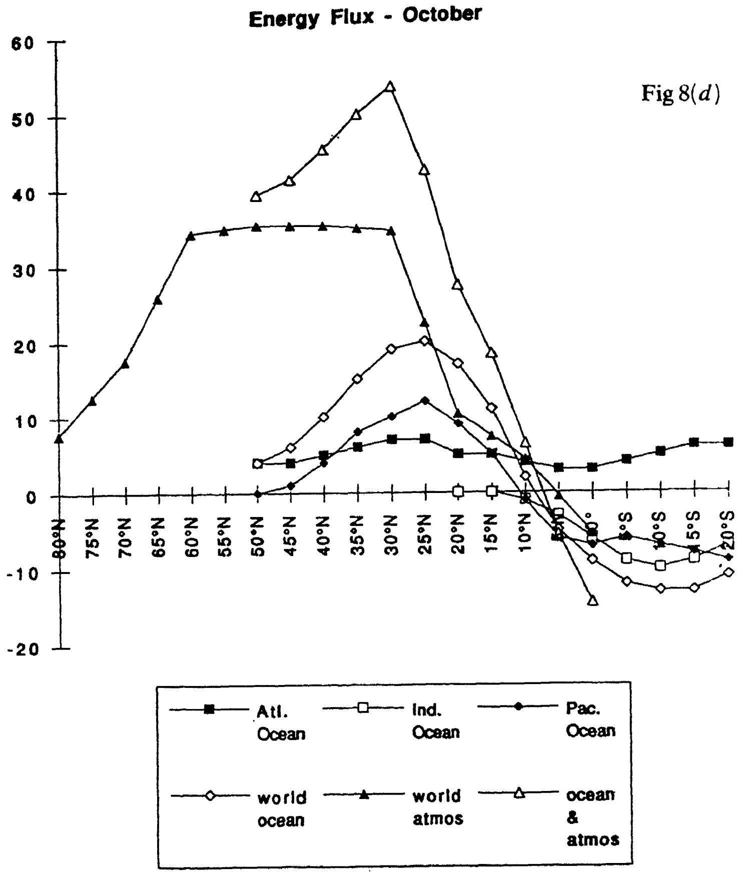


Fig 8 Individual components as indicated of meridional energy flux (units 10^{14} W): (*d*) October

reduction in oceanic flux and the overall energy budget returns to present values. This would have happened at about 18,000-14,000 years BP at the end of the last ice age and is the time at which the oceanic conveyor belt is thought to have "switched on"[45]. It is easy to see that there could have been a net meridional energy flux surplus at this time, quite apart from any Milankovitch variations, so that ice melting at high latitudes would occur. The presence of the ice maintains this steep meridional temperature gradient for some time. The value we deduced from the ice volume change observed between 18000 years BP and the present was about 2.5×10^{14}W—a relatively small number when compared with present day values given in Fig 7*c*.

Conclusions

The issue of temperature limits in various parts of the climate system has been discussed.

The tropical west Pacific warm pool is kept near to 29-30°C by what may be termed evaporational buffering. It is basically the Clausius Clapeyron relationship that governs this limit; the appropriate sensitivity is close to 30 $\mathrm{Wm^{-2}{}^{\circ}C^{-1}}$.

The tropical east Pacific has the same upper limit. The minimum temperatures there, it is suggested, are controlled as a balance between upwelling and associated Ekman divergence, surface heat flux and geostrophic motion towards the equator giving horizontal advection and being controlled by the westward wind stress.

Tropical free troposphere temperatures are thought to be limited by extra water vapor infrared cooling balancing extra near infrared heating, extra condensation and associated adiabatic subsidence, all accompanying El Niño.

Tropical lower stratospheric temperatures rise during volcanic activity. Heating by aerosol absorption is offset by infrared emission from CO_2 which increases at the higher temperatures so that maximum rises appear to be about 6°C.

Ice age minimum temperatures are limited by meridional energy transport by the coupled atmosphere and ocean.

These issues represent a selection of ideas suggesting that temperature limits occur at various points in the climate system and are consistent with physical principles that must be quantitatively understood before we can claim a first order understanding of climate.

Acknowledgements

This work was supported by the National Science Foundation Climate Dynamics Program under Grant No. ATM-91096902 and the National Oceanic and Atmospheric Administration under Grant NA26CP 0021-01. The views expressed herein are those of the authors and do not necessarily reflect the views of NSF or of NOAA or any of its sub-agencies. We thank M Lozano for access to her radiative transfer calculations, Debra Cochrane for her work on Fig. 7 and 8 and on other aspects of the paper and Dorothy Frank for her word processing work on the manuscript over a period of several months.

References

1 J Hsiung *J geophys Res* **91** (1986) 10585-10606
2 R Z Payne *J atmos Sci* **29** (1972) 959-970
3 M I Budyko *Atlas of the Heat Balance of the Earth* Moscow: Gidrometeorozdat (1963) 5 + 69 pp + plates
4 H J Isemer and L Hasse *The Bunker Climate Atlas of the North Atlantic Ocean*, Vol. 2. *Air-sea Interactions* Springer-Verlag Berlin Heidelberg New York Tokyo 24pp with 207 charts and 37 figs (1987)
5 W Large and S Pond *J phys Oceanogr* **12** (1982) 464-482
6 M Bottomley, C K Folland, J J Hsiung, R E Newell and D E Parker *Global Surface Temperature Atlas "GOSTA"* U K Meteorol Off Bracknell (1990) 20 pp and 313 plates
7 R E Newell, A R Navato and J Hsiung *Pageoph* **116** (1978) 351-371
8 R E Newell *Am Scient* **67** (1979) 405-416
9 R E Newell and T G Dopplick *J appl Met* **18** (1979) 822-825
10 R E Newell *J Phys Oceanogr* **16** (1986) 1338-1342
11 J S Godfrey and A C M Beljaars *J Geophys Res* **96** (1991) 22043-22048
12 R E Newell, R Selkirk and W Ebisuzaki *J Climatol* **2** (1982) 781-800

13 R E Newell and Hsiung In: *Climatic Changes on a Yearly to Millennial Basis* (*Eds* N A Morner and W Karlen) Dordrecht: D Reidel Publ Co (1984) 533-561
14 K Wyrtki *J phys Oceanogr* **11** (1981) 1205-1214
15 N E Graham and T P Barnett *Science* **238** (1987) 657-659
16 V Ramanathan and W Collins *Nature* **351** (1991) 27-32
17 J M Wallace *Nature* **357** (1992) 230-231
18 V Ramanathan and W Collins *Nature* **357** (1992) 649
19 D E Waliser and N E Graham *J Geophys Res* **98** (1993) 12881-12893
20 R Fu, A D DeGenio, W B Rossow and W T Liu *Nature* **358** (1992) 394-397
21 R E Newell and Z X Wu *J geophys Res* **97** (1992) 3693-3709
22 R W Spencer and J R Christy *Science* **247** (1990) 1558-1562
23 R E Newell and B C Weare *Science* **194** (1976) 1413-1414
24 J K Angell *J Geophys Res* **93** (1988) 3697-3704
25 A E Gill *Quart J Roy Meteor Soc* **106** (1980) 447-462
26 G H Doherty and R E Newell *Tellus* **36B** (1984) 149-162
27 R N Hoffman *Sci Rep* No. 6 Cambridge MA: Dept Meteor Phys Oceanogr M.I.T. (1981) 124 pp
28 E M Rasmusson In: *The General Circulation of the Tropical Atmosphere* (*Eds* R E Newell *et al.*) M.I.T. Press Cambridge MA (1972) 5, p 215
29 R E Newell *J atmos Sci* **27** (1970) 977-978
30 R E Newell *IRS* '84: *Current Problems in Atmospheric Radiation, Proc int Radiation Symp* (*Ed.* G. Fiocco Hampton) VA: A. Deepak Publishing (1984) 93-101
31 K Labitzke, B Naujokat and M P McCormick *Geophys Res Lett* **10** (1983) 24-26
32 R E Newell and Z X Wu *Bull U S Geol Survey* (1993) (*in press*)
33 CLIMAP Project Members *Science* **191** (1976) 1131-1137
34 *CLIMAP Project Members Seasonal Reconstructions of the Earth's Surface at the Last Glacial Maximum Geol Soc Am Map and Chart Sr* **MC-36** (1981) 18 pp plus maps
35 J A Coetzee *Palaeocology of Africa* **III** (1967) Cape Town/Amsterdam: A A Balkema (1967) 146 pp
36 T Van der Hammen *J Biogeogr* **1** (1974) 3-26
37 R E Newell, C F Herman, S Gould-Stewart and M Tanaka *Nature* **253** (1975) 33-34
38 R F Flint *Glacial and Quaternary Geology* John Wiley and Sons Inc NY, London (1971) 417-421
39 W L Gates *J atmos Sci* **33** (1976) 1844-1873
40 R E Newell, S Gould-Stewart and J C Chung In: *Palaeoecology of Africa* **13** (*Eds* J A Coetzee and E M Van Zinderen Bakker) A Balkema Rotterdam (1981) 1-19
41 R E Newell *Quart Res* **4** (1974) 117-127
42 A H Oort and T H Vonder Haar *J Phys Oceanogr* **6** (1976) 781-800
43 J Hsiung, R E Newell and T Houghtby *Quart J R meteor Soc* **115** (1989) 1-28
44 A L Gordon *J geophys Res* **91** (1986) 5037-5046
45 W S Broecker and G H Denton *Geochim Cosmochim Acta* **53** (1989) 2465-2501

A STUDY ON THE DYNAMICS AND ENERGETICS OF THE INDIAN SUMMER MONSOON

U C MOHANTY* AND K J RAMESH

Centre for Atmospheric Sciences, Indian Institute of Technology
Hauz Khas, New Delhi-110 016, *India*

(*Received* 12 *July* 1993; *Accepted* 30 *August* 1993)

The present study is carried out to study the mean circulation characteristics, the dynamics and energetics of the Indian summer monsoon through the analysis of kinetic energy, vorticity, angular momentum, heat and moisture budgets with the help of long-term mean (1979-88) analyzed data sets. The data base used in this study consists of the uninitialized analyses (12 Z) of the European Centre for Medium Range Weather Forecasts (ECMWF) for ten years (1979-88). In this set of ten years data base, the data for the 1979 summer is the part of the FGGE level IIIb analysis and the remaining data of nine summer seasons comprise the global analyses of the ECMWF.

It is found that the summer monsoon circulation constitutes a major centre of low level convergence and upper level divergence over South East Asia. The circulation over the East African continent and the adjoining western Indian Ocean determines the strength of the ageostrophic motions in maintaining the lower tropospheric features of the monsoon. The entrance (exit) regions of the tropical easterly jet (TEJ) are characterized by production (destruction) of kinetic energy which is essential to maintain the outflow (inflow) prevailing at the respective locations of the TEJ. Further, the study illustrates the role of moisture convergence and the release of diabatic heating to maintain the entire quasi-permanent monsoon circulation over South East Asia.

Key Words: Indian Summer Monsoon; Circulation Characteristics; Dynamics; Energetics

Introduction

The Indian summer monsoon circulation represents some of the most prominent changes of the atmospheric circulation such as periodic movements in the locations of jet streams and reversal of wind regimes which occur with surprising regularity every year and this circulation is basically generated due to the thermal contrasts (differential heating) between the land and the ocean as a consequence of the difference in their response to solar radiation. After the completion of several field experiments such as International Indian Ocean Experiment (IIOE), Indo-Soviet Monsoon Experiment (ISMEX), MONSOON-77 and MONEX-79, it is well recognized that a better understanding of the dynamics and energetics of the summer monsoon through the data based diagnostic studies is very much essential for improved simulation/prediction. The prominence of the diagnostic studies with the observed data sets is very well known (Pearce[1], Kung and Smith[2]) as an important GARP problem in understanding the tropical dynamics.

**Present Affiliation*: National Centre for Medium Range Weather Forecasting (NCMRWF), INSAT Building, Mausam Bhavan Complex, Lodi Road, New Delhi-110 003.

Though, a large number of studies are carried out to understand the atmospheric dynamics, most of them are confined to the data rich middle latitude regions such as North America and British Isles. Many of the studies are mainly conducted for individual middle latitude systems. Only few studies are found over the tropics, in particular, over the Indian summer monsoon (Kung[3,4]; Chen[5]; Oort and Chan[6]; Anjaneyulu[7]; Ananthakrishnan and Keshavamurty[8]; Keshavamurty and Awade[9]; Singh *et al.*[10]; Mandal *et al.*[11]—through the analysis of kinetic energy budget; Holton and Colton[12]; Fein[13]; Chu *et al.*[14]; Daggupaty and Sikka[15]—through the analysis of vorticity budget; Rao[16]; Sankar Rao and Ramanatham[17]; Keshavamurty[18]; Oort and Chan[6] and Newton[19,20]—through the analysis of angular momentum budget; Mohanty *et al.*[21,22,23]; Pearce and Mohanty[24]—through the analysis of heat and moisture budgets). Most of them are either case studies using special data sets or studies restricted to only certain terms of the large scale budgets. First GARP Global Experiment provided an opportunity to undertake detailed studies over the Indian summer monsoon[21,22,23,24].

In this paper, an attempt is made to study the circulation characteristics, dynamics and energetics of the Indian summer monsoon with the help of the long-term mean analyzed data sets.

Methodology

In this study to understand the dynamics and energetics of the Indian summer monsoon circulation, all the basic equations which form a closed system in describing the atmospheric circulation involving equations of motion, continuity equation, thermodynamic energy equation and moisture continuity equation are considered. The same set of equations are also used in the formulation of various global/regional atmospheric models for simulation of the atmospheric circulation and the prediction of weather systems. A detailed analysis of the mean dynamical characteristics of the summer monsoon is carried out through the computation of kinetic energy, vorticity and angular momentum budgets of the monsoon regime. Further, the mean thermodynamical characteristics of the monsoon circulation are studied through the analysis of the heat and moisture budgets during the summer monsoon season (JJA). These budget equations are derived from the closed system of atmospheric model through simple mathematical treatments.

The following are the various budget equations used in the present study and all of them are expressed in flux form involving pressure coordinates in the vertical:

Mean Kinetic Energy Budget

$$\frac{\partial K_M}{\partial t} + (\nabla \cdot H_0 + \nabla \cdot H_1) + \partial\overline{(K_M + \overline{V} V')\,\omega}/\partial p = -\,\overline{V} \cdot \nabla \overline{\Phi} - C(K_M, K_T) - \overline{V} \cdot \overline{F}, \quad \ldots (1)$$

where

$K_M = 1/2\ \overline{V}^2$,

$$H_0 = K_M \overline{V}$$

$$H_1 = (V \cdot \overline{V') V'},$$

$$C(K_M, K_T) = C_H(K_M, K_T) + C_V(K_M, K_T),$$

$$C_H(K_M, K_T) = -\overline{u'u'}/a\cos\phi \frac{u'v'\cos\phi}{} \frac{\partial(\bar{u}/a\cos\phi)}{}$$

$$\overline{u'v'}/a\cos\phi \frac{\partial\bar{v}}{\partial\lambda} - \frac{\overline{v'v'}}{a}\frac{\partial\bar{v}}{\partial\phi}$$

$$u'u' \quad \bar{v}\tan\phi$$

and

$$C_v(K_M, K_T) = -\frac{a}{u'\omega'}\frac{\partial\bar{u}}{\partial p} - v'\omega' \frac{\partial\bar{v}}{\partial p}.$$

The first term on the L.H.S. signifies the local rate of change of kinetic energy. The second term describes the horizontal flux divergence and the third term describes the vertical flux divergence of kinetic energy. Similarly, the first term on the R.H.S. denotes the adiabatic production of kinetic energy and second term describes the exchange of energy between the mean flow and the transients. The last term denotes the dissipation of kinetic energy by the turbulent frictional processes.

Mean Vorticity Budget Equation

$$\frac{\overline{\partial\zeta}}{\partial t} + \nabla\cdot(\overline{\zeta}+f)\,\overline{V} + \frac{\partial(\overline{\zeta}\bar{\omega})}{\partial p} = -(\overline{\zeta}+f)\,D - k\cdot\nabla\omega\times\frac{\partial V}{\partial p} + Z, \qquad \dots (2)$$

where ζ is the relative vorticity, f is the coriolis parameter, D is the horizontal divergence and Z is the vorticity residual usually considered as an apparent vorticity source/sink. As explained earlier, first term on the L.H.S. signifies the local rate of change of vorticity. The second and third terms represent horizontal flux divergence and the vertical divergence of vorticity respectively. The first and second terms on the R.H.S. describe the production of vorticity through the stretching of vortex tubes and the tilting of isobaric surfaces respectively.

Mean Angular Momentum Budget

$$\frac{\overline{\partial M}}{\partial t} + \nabla\cdot\overline{M}\overline{V} + \frac{\partial(\overline{M}\bar{\omega})}{\partial p} - f\,\overline{V}a\cos\phi = -\frac{\partial\overline{\Phi}}{\partial\lambda} + \delta\Phi + F_\lambda, \qquad \dots (3)$$

where $M = u\,a\cos\phi$ is the zonal momentum and Φ is the geopotential. The principle of conservation of angular momentum arises from Newton's second law of motion and states that the angular momentum of a system will change only if external torques are applied. Atmosphere is subjected to large torques exerted through its interaction with the surface of the earth. The first term of the L.H.S. signifies the local rate of change of zonal momentum flux, the second and third terms describe the horizontal and vertical flux divergence of zonal momentum flux by the mean flow. The fourth term denotes the Ω-momentum flux divergence. The first term on the R.H.S. represents the production/destruction of momentum due to zonal (east-west) pressure gradients (pressure torque) and the second term describes the mountain torque (basically due to the presence of mountains). The last term represents the frictional torque.

Mean Heat and Moisture Budget Equations

$$\overline{\frac{\partial(C_P T)}{\partial t}} + \nabla\cdot(C_P\overline{T}\overline{V}) + \genfrac{}{}{0pt}{}{\partial(C_P\overline{T}\bar{\omega})}{\partial p} \quad \bar{\omega}\,\bar{\alpha} = QS \qquad \ldots (4)$$

and

$$\overline{\frac{\partial(Lq)}{\partial t}} + \nabla\cdot(L\bar{q}\overline{V}) + \frac{\partial(L\bar{q}\bar{\omega})}{\partial p} = QL\,, \qquad \ldots (5)$$

where C_P is the specific heat at constant pressure, L is the latent heat of evaporation, q is the specific humidity and T is the temperature. The term QS represents the net effect of all diabatic contributions consisting mainly of radiative effects, latent heat release due to condensation and reevaporation of water vapor of falling precipitation and turbulent heat transfer. The term QL represents the net effect of all diabatic contributions to the latent heat basically due to the turbulent transfer and effects of condensation and evaporation. The remaining terms of the equations have their usual meaning.

In this study all the dynamical and thermodynamical budgets are computed using ten year mean analyzed fields for the summer monsoon season (JJA) and hence all the terms representing local time rate of change of various quantities have negligible contribution towards the respective budgets. The results at each regular latitude-longitude grid points of 1.875° are averaged both in zonal and meridional directions over the summer monsoon region and integrated vertically from 1000hpa to 100hpa layer. Thus, the volume integration of any variable $F(\lambda, \phi, p)$ for a limited region bounded by meridians λ_1 and λ_2, latitude circles ϕ_1 and ϕ_2 and isobaric surfaces p_1 and p_2 may be written as

$$\overline{F} = \frac{1}{g}\int_{\lambda}^{\lambda_2}\int_{\phi_1}^{\phi_2}\int_{p_1}^{p_2} F\,a^2\cos\phi\;\partial\lambda\;\partial\phi\;\partial p\,,$$

where a is the average radius of the earth. The volume integration of all the budget equations with the boundary conditions that ω vanishes at the bottom and the top of the atmosphere leads to the elimination of all the terms representing vertical flux divergence of various quantities. All the last terms in the R.H.S. of the budget equations (1-5) represent the contributions from the sub-grid scale processes. These terms are evaluated implicitly in this study as a residue of all the other terms in the respective budget equations. In this study, all the terms in the various budget equations mentioned above are evaluated over the summer monsoon region [30°E-150°E; 15°S-30°N]. Though, all the budget terms in equations 1-5 are estimated the discussions in this paper are restricted to those budget terms which contribute significantly to the dynamical and thermodynamical budget equations and seems to be responsible to the maintenance of the summer monsoon circulation.

Data Set and Analysis Procedure

The data base used in this study consists of the uninitialized analyses (12*Z*) of the European Centre for Medium Range Weather Forecasts (ECMWF) for ten years (1979-88). In this set of ten years data base, the data for the 1979 summer is the part of the FGGE level IIIb analysis and the remaining data of nine summer seasons comprise the global analyses of the ECMWF. The basic meteorological fields considered for the study include wind, temperature, geopotential and moisture over the summer monsoon region extending from 45°N to 45°S and 30°E to 150°E at ten pressure levels (1000, 850, 700, 500, 400, 300, 200, 150, 100 and 50hpa). The mass and velocity fields are analyzed by a four dimensional multivariate data assimilation scheme developed and processed at the ECMWF. Details of the data assimilation are given by Lorenc[25], Bengtsson *et al.*[26] and Lonnberg and Shaw[27]. In this scheme, the assimilation is carried out in three different steps viz., multivariate optimal interpolation, nonlinear normal mode initialization and six hour forecast which provides the first guess for the subsequent analysis.

The initialized vertical velocity field obtained from the adiabatic/diabatic non-linear normal mode initialization and stored in the ECMWF archives does not represent true vertical motion in the tropics (Kanamitsu[28]; Mohanty *et al.*[23]). Further, the changes incorporated in the analysis procedure from time to time at the ECMWF have had the greater impact on divergent winds and vertical velocity (Hoskins *et al.*[29]). It was in fact suggested by Julian[30] that the horizontal wind fields generated by the ECMWF analysis represent a reasonable planetary scale divergent circulation over the tropics. Hence, in this study, instead of using the ω-field stored in ECMWF data archives, ω-field is derived through the kinematic technique using uninitialized wind fields is used. In the kinematic method, the main problem is that of the accumulated biased errors involved in the computation of divergence from the wind components. A technique suggested by O'Brien[31] is used to adjust the divergence in such a way that its vertically integrated value in any particular column of the atmosphere becomes zero. The consistency of the vertical motion thus computed is found to represent better semi-permanent circulation (Hadley cell) in the tropics (Mohanty *et al.*[32]). The vertical velocity fields computed by the kinematic method are used in the present study for the analysis of dynamics and energetics of the Indian summer monsoon.

All the terms described in the previous section are calculated using the ECMWF analyses for the ten summer monsoon seasons (June-August) from 1979 to 1988. The various space derivatives are evaluated by centred difference scheme. All the budget terms are computed at each regular latitude-longitude grid points and then integrated suitably over the domain of the atmosphere to obtain vertically integrated geographical distributions, sectorial mean vertical cross-sections, latitudinal distributions of sectorially averaged and vertical integrated fields and volume integrated fields of various budget parameters.

Results and Discussion

The analysis of some of the large scale meteorological fields and significant budget terms is presented to illustrate the circulation characteristics, dynamics and energetics of the Asian summer monsoon.

Monsoon Circulation Characteristics

The mean wind field at 850hpa and 150hpa are presented in figures 1(*a*) & 1(*b*) respectively. In general, the significant monsoon circulation features of these two pressure levels such as the low level Somali jet and the upper level tropical easterly jet (TEJ) are in well agreement with their respective climatological features presented by Newell *et al.*[33]; Rao[34], and Das[35]. The distribution of mean wind field at 850hpa shows the existence of strong southern hemispheric trades having a speed of about 10ms^{-1} (50°E-80°E), strong cross-equatorial flow into the Northern Hemisphere off the Somali Coast and a strong zone of westerlies over the Arabian Sea and the Bay of Bengal (10-20ms^{-1}). Wind field distribution at 150hpa is characterized by a strong and elongated anti-cyclone centered over Tibet (30°N) in the Northern Hemisphere. Two contrasting wind regimes present on either side of the Tibetan anti-cyclone i.e., westerly wind regime (sub-tropical westerly jet) towards the north and easterly regime (tropical easterly jet) to the south are the most significant circulation features of the Indian summer monsoon. Zone of easterlies extends over Indonesia, Bay of Bengal, India, Arabian Sea and Africa with a core of maximum (>20ms^{-1}) at around 10°-12°N over the south Indian peninsula and adjoining Arabian Sea. The westerly wind flow over the sub-tropical regions of the Southern Hemisphere is stronger than the northern hemispheric westerly regime as the winter hemispheric westerly jet is always stronger.

Geographical distribution of geopotential height at 850hpa (Fig. 1*c*) illustrate the presence of a low over the North West India and adjoining Pakistan with an elongated monsoon trough over the North India. The formation and existence of this low over the South Asian land mass is largely attributed to the strong horizontal gradients of geopotential generated due to the differential heating of the land and the oceans and is known as heat low. The subtropical highs over the Southern Hemisphere surrounding the southern African continent and the adjoining Indian Ocean sectors are responsible for a strong easterly trade winds over the southern hemispheric equatorial tropics and strong cross-equatorial flow off the east coast of Africa (Fig. 1*a*). The low of

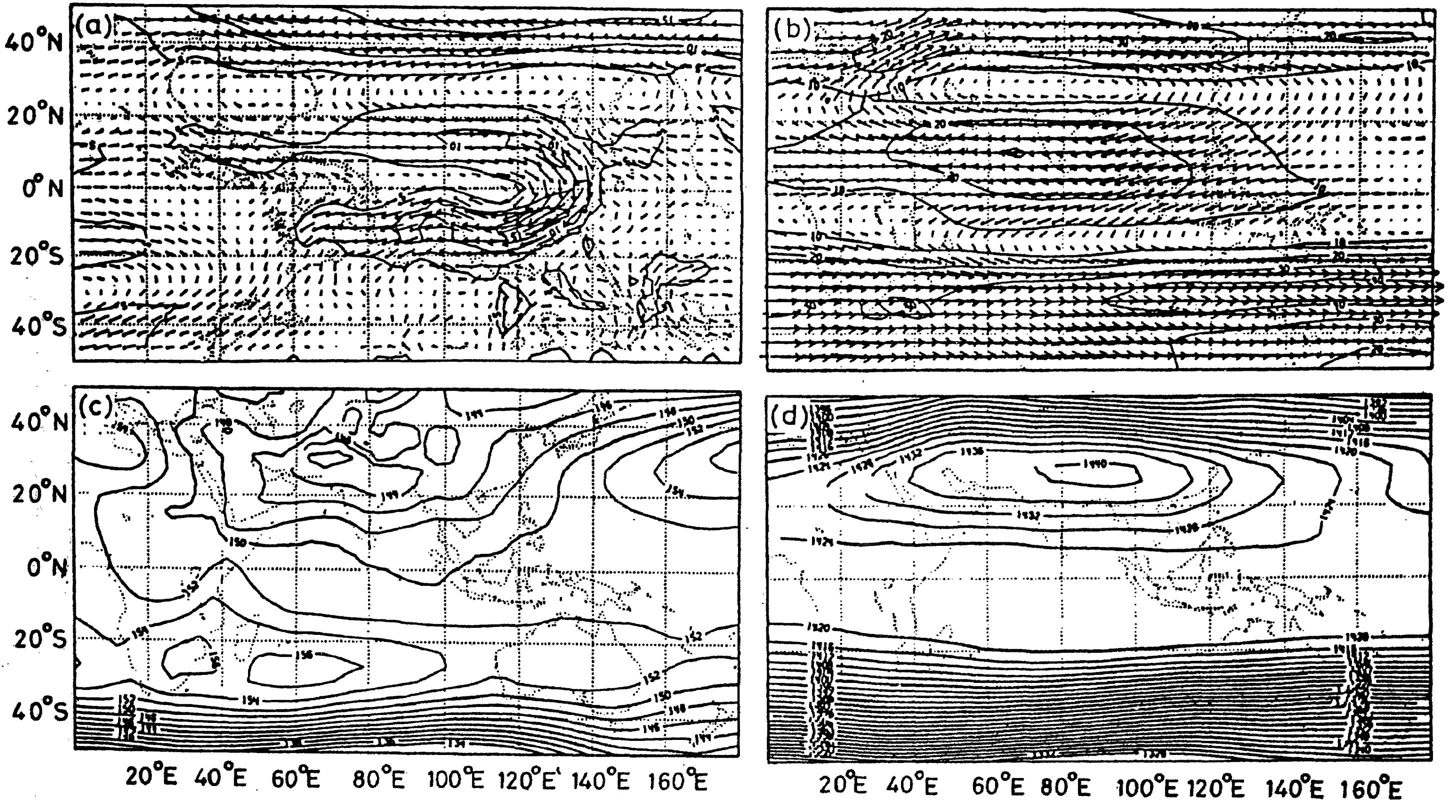


Fig 1 Geographical distribution of mean wind (V) and geopotential (ϕ), (*a*) V850hpa (ms^{-1}), contour interval: $5ms^{-1}$; (*b*) V at 150hpa (ms^{-1}), interval $10ms^{-1}$; (*c*) ϕ at 850pa (gpm), contour interval: 2×10^1gpm; (*d*) at 150hpa (gpm), contour interval: 4×10^1gpm

geopotential found over the South Asian land mass is flanked by geopotential height situated over the Mediterranean and the North Pacific Oceans respectively. The upper level geopotential height distribution (Fig. 1*d*) is characterized by an intense and elongated high pressure cell over the Tibet. This semi-permanent anti-cyclone over Tibet is responsible for the prevailing easterly regime over the Indian peninsula with the formation of the tropical easterly jet (TEJ).

The summer monsoon circulation is generally visualized as a large scale convergence of mass and moisture over the Indian sub-continent and adjoining South East Asian land mass in the lower levels and a strong upper level divergence aloft. In order to examine the characteristics of the divergent circulation associated with the monsoon, the velocity potential and the divergent wind components at 850hpa and 150hpa are presented in Fig. 2. At 850hpa around 20°S, a strong divergence center (χ-minimum) is located off the east coast of South Africa and Madagascar (Figs 2*a* & *b*) which is considered to control the strength of the cross-equatorial current and moisture transport from the Southern Hemisphere to the monsoon region. The major centre of low level convergence located over the North East India, Burma and adjoining land mass is the most active zone with copious rainfall during the monsoon season. An exactly opposite circulation feature is observed in the upper level at 150hpa with a strong center of divergence (χ-minimum) over the South East Asia and the western Pacific and a strong convergence regime over the Indian Ocean to the south of the equator. The presence of strong east-west circulation (having large velocity potential gradients) are the characteristic phenomena pertaining to the summer monsoon (Krishnamurti *et al.*[36]). These circulations in a zonal plane having a strong warm air ascent over the South East Asia and descent over the Central/Eastern Pacific Ocean are important for the generation of kinetic energy of the quasi-stationary ultra long waves. On the other hand, the north-south circulation is important for the generation of kinetic energy of the mean zonal flows. The presence of Tibetan anti-cyclone with its strong divergent circulation is essential for the maintenance of both the circulations in zonal and meridional planes. Hence, a very strong upper level divergent circulation is also crucial for the maintenance of the summer monsoon.

In order to examine the dominant characteristics of the monsoon circulation in the vertical plane and also to demonstrate the area of convergence/divergence and vertical motion, the sectorial mean (30°E-150°E) vertical cross-sections of zonal wind (u), meridional wind (v), horizontal divergence (D) and the vertical motion (ω) are presented in Fig. 3. The cross-section of the zonal wind averaged over the monsoon region (30°E-150°E) agrees very well with the general pattern of two westerly wind maxima separated by an equatorial maximum of easterly wind in the upper tropospheric levels (Fig. 3*a*). Further, the presence of monsoon westerlies in the tropical lower troposphere up to 500hpa of the Northern Hemisphere and easterlies in the Southern Hemisphere constitute the basic flow pattern during the summer (JJA).

The cross-section of the meridional wind (Fig. 3*b*) is characterized by a strong cross-equatorial flow both in the lower and upper tropospheric levels with a flow across the equator into the Northern Hemisphere in the lower le-

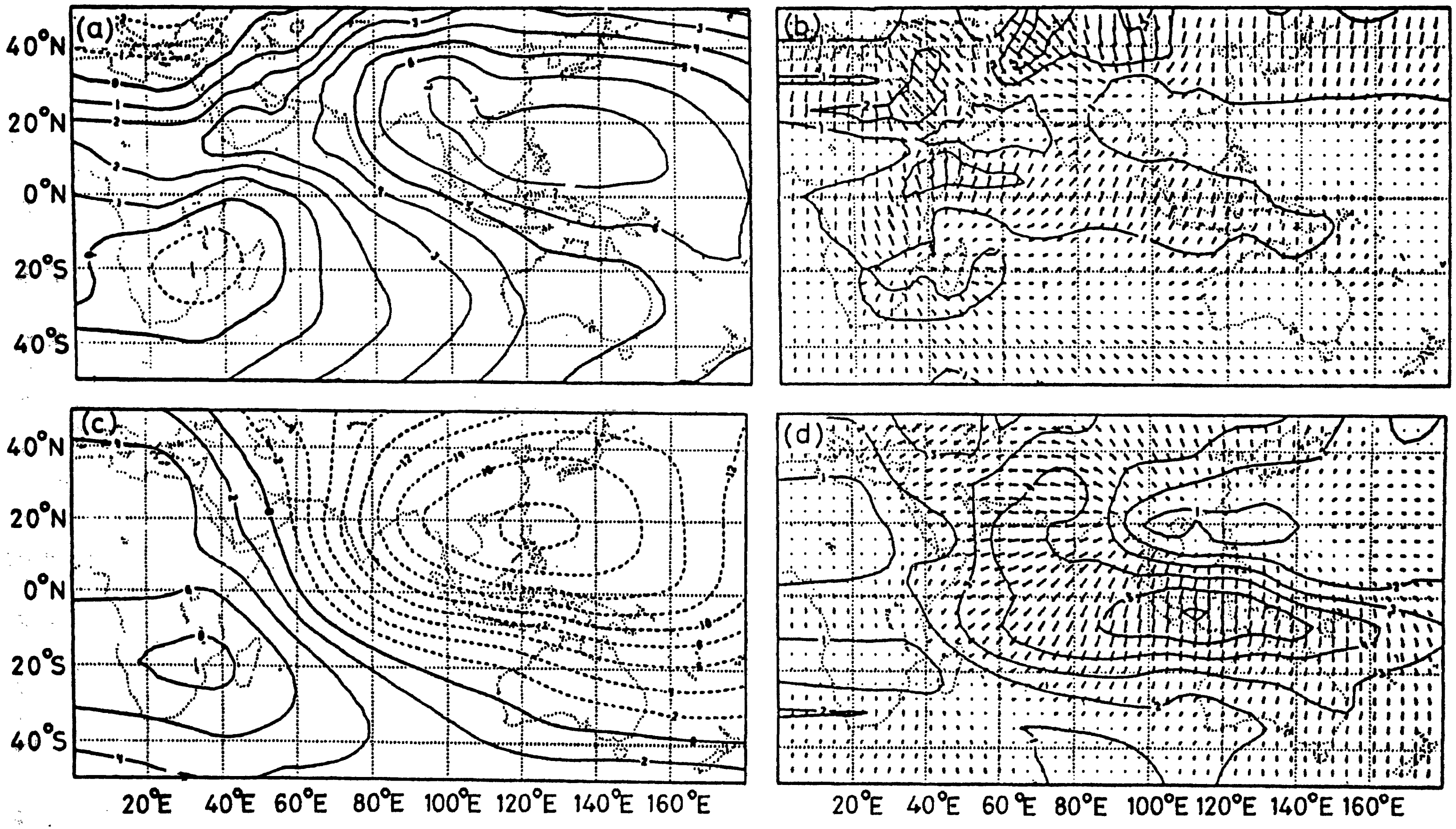


Fig 2 Geographical distribution of velocity potential (χ) and divergent wind (Vχ); (*a*) χ at 850hpa ($10^5 m^2 s^{-1}$) Contour interval: 1 × 10; (*b*) Vχ at 850hpa ms^{-1}, contour interval: 1, (*c*) same as (*a*) but for 150hpa contour interval: 2 × 10; (*d*) same as (*b*) but for 150hpa.

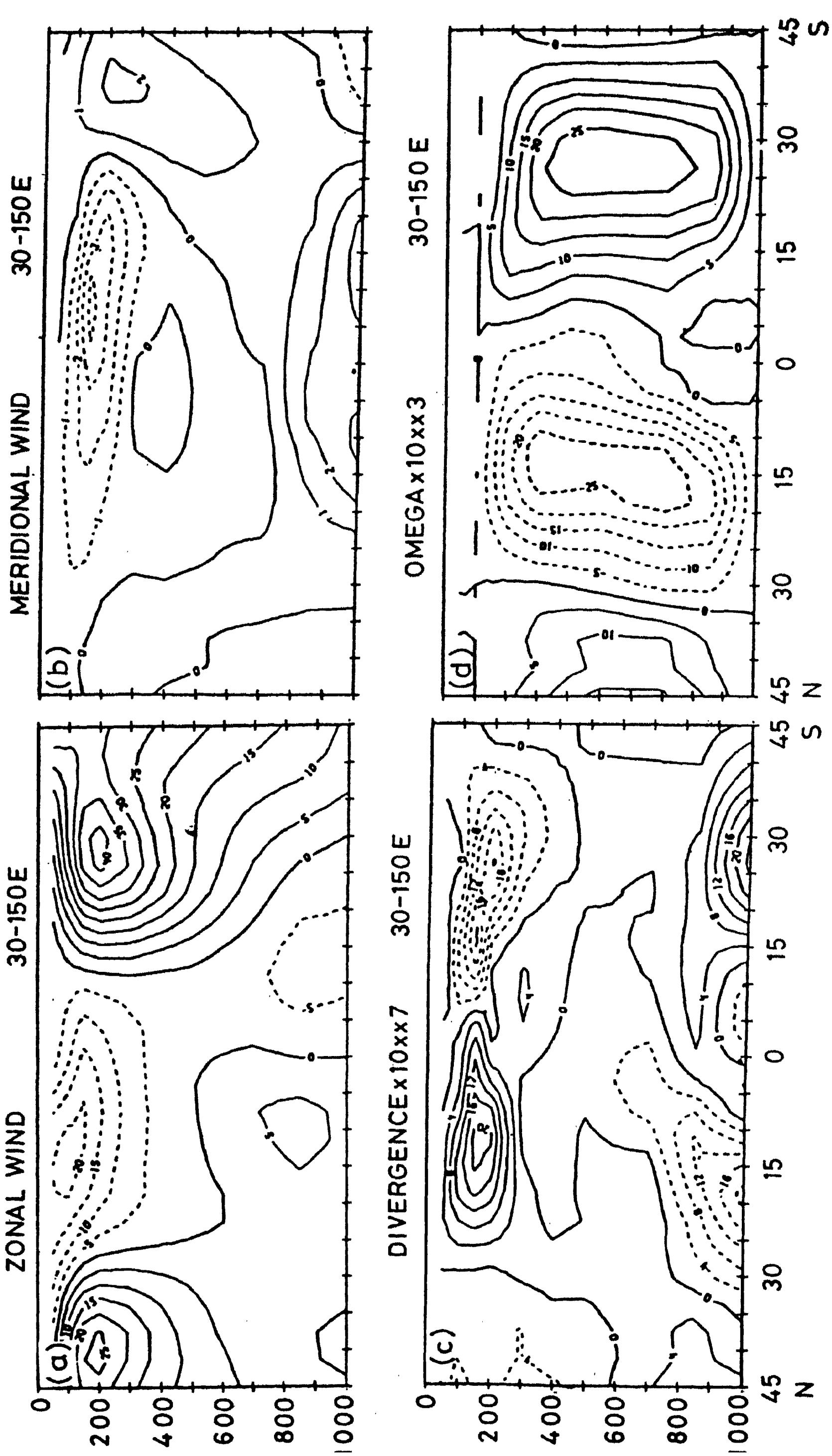


Fig 3 Sectorial mean (30°E-150°E) pressure-latitude cross-sections of (*a*) zonal wind ms^{-1}, contour interval: 5ms^{-1}; (*b*) meridional wind ms^{-1}, contour interval: 1ms^{-1}; (*c*) divergence (10^{-7} sec^{-1}, contour interval 4; (*d*) vertical velocity 10^{-3} Pa sec^{-1}, contour interval: 5

vels (1000-700hpa) and a strong reverse flow in the upper levels (300-100hpa). While the lower tropospheric southerly flow is originated from the sub-tropical high pressure cells and divergent centers of the Southern Hemisphere, the upper level strong northerly flow prevails mainly due to the strong outflow emanating from the Tibetan anti-cyclone.

Distribution of the divergence field (Fig. 3*c*) displays a deep layer of low level convergence and a strong and norrow layer of upper level divergence over the monsoon latitudes (10°-30°N). An exactly opposite circulation with strong upper level convergence and low level divergence is dominated in the southern hemispheric tropics. These characteristics of the divergence field are responsible for the maintenance of the prevailing zone of rising motion in the Northern Hemisphere and the zone of subsidence (downward motion) in the Southern Hemisphere (Fig. 3*d*) during the northern summer (JJA).

Thus, the circulation characteristics of the Indian summer monsoon based on the ECMWF analyses of 10-years (1979-88) represent the major circulation features of the summer monsoon climatology and quite reliable for undertaking diagnostic studies.

To examine the basic thermodynamic parameters of the monsoon, the horizontal distributions of the mean tropospheric temperature and net tropospheric moisture together with their zonal asymmetric anomalies (computed from the global zonal mean values) are illustrated in Fig. 4. Mean tropospheric temperature computed from the 10 years of the ECMWF analyses for JJA (Fig. 4*a*) shows that the temperature maximum is situated over the south Asian region covering mainly the Arabia, the Afghanistan and the Indian sub-continent. This zone of maximum temperature over the South Asia plays a very important role in the build-up of necessary land-ocean temperature contrast and the north-south pressure gradients for the establishment and maintenance of the cross-equatorial flow and the summer monsoon circulation over the South Asia. Even the zonal asymmetric mean tropospheric temperature analysis (Fig. 4*b*) also confirms the presence of a strong zonal anomaly of temperature to an extent of 5°C over the south west Asiatic land mass.

Distribution of the vertically integrated (1000-300hpa) net tropospheric moisture (Fig. 4*c*) which also represents total precipitable water content for the summer season (JJA) indicates three zones of humidity maxima. Among them, the primary maximum of humidity is situated over the North East India and others are located over the Ethiopia and west central Pacific Ocean respectively. Zonal asymmetric anomaly maxima (Fig. 4*d*) are also situated over the same locations as that of net tropospheric moisture maxima. Thus, the area under the influence of the summer monsoon is characterized by a mean tropospheric warming together with a maximum moisture build up for the development of intense convective activity and copious rainfall surrounding the Indian sub-continent during the onset and maintenance of the summer monsoon regime.

Kinetic Energy Budget

The local balance of kinetic energy (KE) is by and large governed by three terms. They are the KE flux divergence, the adiabatic generation of KE and

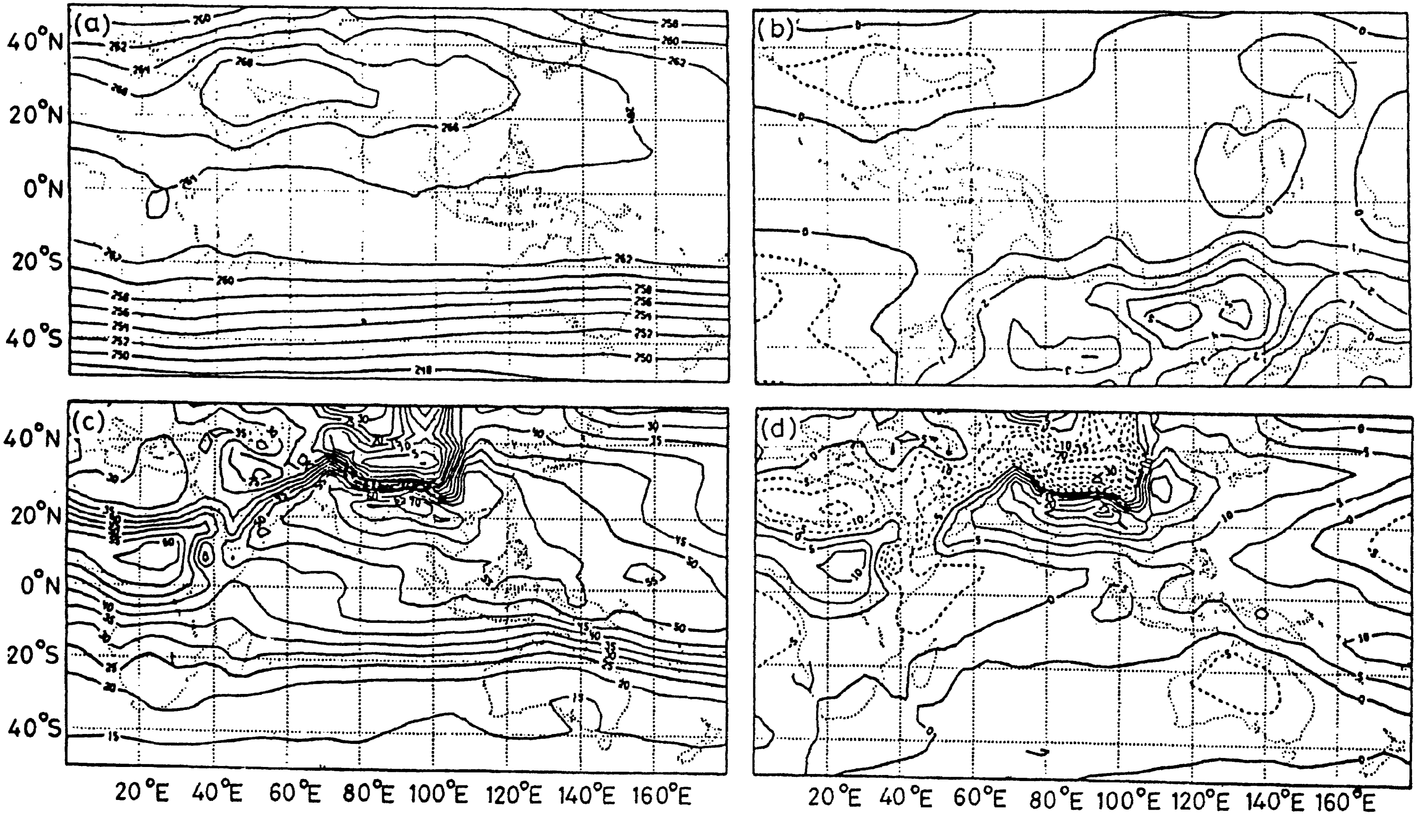


Fig 4 Geographical distribution of (*a*) Mean tropospheric temperature K, contour interval: 2; (*b*) Zonal asymmetric temperature c, contour interval: 1; (*c*) Net tropospheric moisture %, contour interval; 5; (*d*) Same as (*c*) but for zonal asymmetric moisture.

the dissipation of KE (sub-grid scale frictional forces). The mean flow kinetic energy flux divergence can be expressed in terms of two components which can be written as

$$\nabla H = \nabla H_0 + \nabla H_1$$

The first term on the R.H.S. represents the flux transport of kinetic energy by the mean flow and the second term represents the mean flow kinetic energy flux transport by the transient eddy flow. The relative magnitudes of the flux transport of kinetic energy by the eddy flow are found to be small in comparison with the mean flow transport (Mohanty *et al.*[32]).

Vertically integrated (1000-100hpa) geographical distribution of some of the significant kinetic energy budget terms over the summer monsoon region (30°N-15°S; 30°E-150°E) are presented in Fig. 5. It is found that a zone of kinetic energy flux divergence is extending all over the south Asian region covering from the western Pacific to eastern Arabian Sea (with flux divergence maxima situated over the Bay of Bengal and the eastern Arabian Sea respectively). The kinetic energy flux convergence is observed over the western Arabian Sea, adjoining Arabia and north African regions. These zones of kinetic energy flux transport maxima/minima are situated at the respective locations of entrance/exit regions of the tropical easterly jet.

Horizontal distribution of kinetic energy production/destruction is depicted in Fig. 5*b*. The kinetic energy is basically produced by the ageostrophic component of the flow. Positive magnitudes signify the generation of kinetic energy from the available potential energy (APE) source and the negative magnitudes denote the destruction of kinetic energy i.e., transformation of kinetic energy back to the APE source. It is found that the areas characterized by the flux divergence of kinetic energy are the regions of strong kinetic energy production and regions of flux convergence are characterized by destruction/weak production of kinetic energy. Such a nature of kinetic energy production is in fact necessary to maintain the strong outflow/inflow of energy at the entrance/exit regions of the TEJ in the upper troposphere.

It is interesting to examine the contributions of the zonal and meridional components of the adiabatic production of kinetic energy (Figs 5*c* & *d*) over the summer monsoon region. It is found in the earlier studies (Kung[3] and others) that both the components of kinetic energy production tend to oppose each other with a general production of kinetic energy by the meridional component and destruction by the zonal component. But over the summer monsoon region, the zonal component of generation is contributing towards the kinetic energy production over the eastern Arabian Sea and adjoining peninsular India (Fig. 5*c*) which is contrary to its general character found elsewhere over the globe. As seen earlier from the analysis of kinetic energy flux divergence and the adiabatic production, the summer monsoon circulation is characterized by two centers of action, one over the Bay of Bengal and another over the eastern Arabian Sea and adjoining South Indian peninsula. The existence of two kinetic energy production maxima over the above mentioned locations is vital for the maintenance of the flux divergence maxima situated over the same locations. It is interesting to note that the eastern Arabian Sea maximum

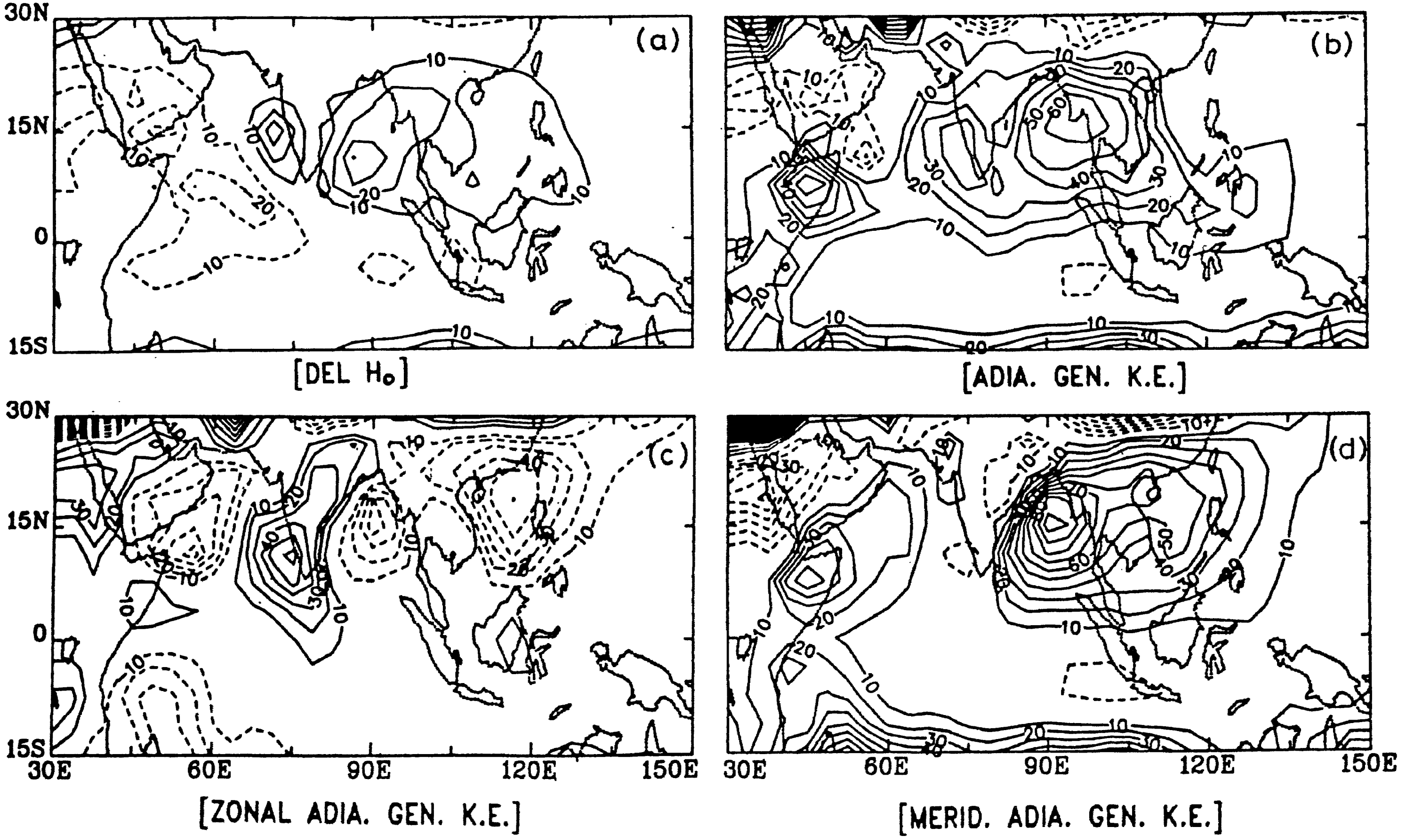


Fig 5 Geographical distribution of vertically integrated kinetic energy budget terms for JJA (1979-88) 10^{-1} Wm^{-2}), contour interval: 10, (*a*) ∇H_0; (*b*) Adiabatic generation of KE; (*c*) Zonal adiabatic generation of KE; (*d*) Meridional generation of KE

of kinetic energy production is maintained by the zonal part of the ageostrophic flow and that of the Bay of Bengal is maintained by the meridional ageostrophic flow of the summer monsoon circulation (Figs 5*c*-*d*). Hence, in the maintenance of the summer monsoon circulation, the adiabatic production of kinetic energy through the action of the pressure forces plays a very important role.

The sectorial mean (30°E-150°E) vertical distribution of significant kinetic energy budget terms for the summer monsoon season (JJA) are illustrated in Fig. 6. The magnitudes of kinetic energy flux transport by the mean flow (Fig. 6*a*) and the transient eddies (Fig 6*b*) are comparable in both the hemispheres during the summer. It is found that lower tropospheric monsoon circulation over the NH-tropics is characterized by a zone of flux divergence over the equatorial region (0°-10°N) and a zone of flux convergence to the north (10°N-22.5°N). In the upper troposphere, the cross-sections of the flux transport terms exhibit very interesting characteristics. Distribution of ∇H_0 shows a zone of kinetic energy flux divergence in the upper troposphere (300-100hpa) in the tropics (7.5°N-26°N) and a zone of flux convergence in the extra tropics. But the role of the Reynold's stresses on the flux transport of kinetic energy ($\nabla \cdot H_1$) is found to be counter balancing the mean flow flux transport over the extra tropics with a strong flux transport of kinetic energy. In the tropical upper troposphere of the summer monsoon regime, the transient eddies are found to supplement the mean flow transport leading to a net flux transport of kinetic energy from the region. In the upper troposphere of the Southern Hemisphere both the kinetic energy flux transport terms contribute equally to the strong flux divergence in the tropics and the strong flux convergence in the extra tropics. In the free atmosphere, the adiabatic production of kinetic energy is found to be large over the regions of strong flux divergence and a destruction/weak production of kinetic energy over the regions of flux convergence (Fig. 6*c*). The variation of the kinetic energy generation illustrate two maxima, the primary one in the boundary layer (1000-850hpa) and the other in the upper troposphere (200-100hpa). Similar results are found in the studies of Kung[3] over the data rich North American continent. Apart from this, strong zones of kinetic energy generation are found in the boundary layers of both the hemispheres. Interestingly, over the Northern Hemisphere, the zone of kinetic energy production in the planetary boundary layer is mainly limited to the tropical latitudes (0°-20°N) indicating the presence of strong ageostrophic (cross-isobaric) flow in the lower troposphere of the monsoon regime. In the Southern Hemisphere, however, the areas of strong kinetic energy production are extended up to the mid-latitudes.

Latitudinal distribution of the vertically integrated and kinetic energy budget terms averaged over the monsoon region (30°E-150°E) are depicted in Fig. 7. In an earlier study by Mohanty *et. al*[32], it was found earlier that the magnitudes of the kinetic energy are very small over the tropical latitudes and are large over the winter hemispherie westerly jet regime. The present study confirms the fact that the magnitudes of KE over the extra tropics of the Southern Hemisphere are quite large (Fig. 7*a*). However, over the NH-tropics, the net tropospheric kinetic energy magnitudes are significant only due to the

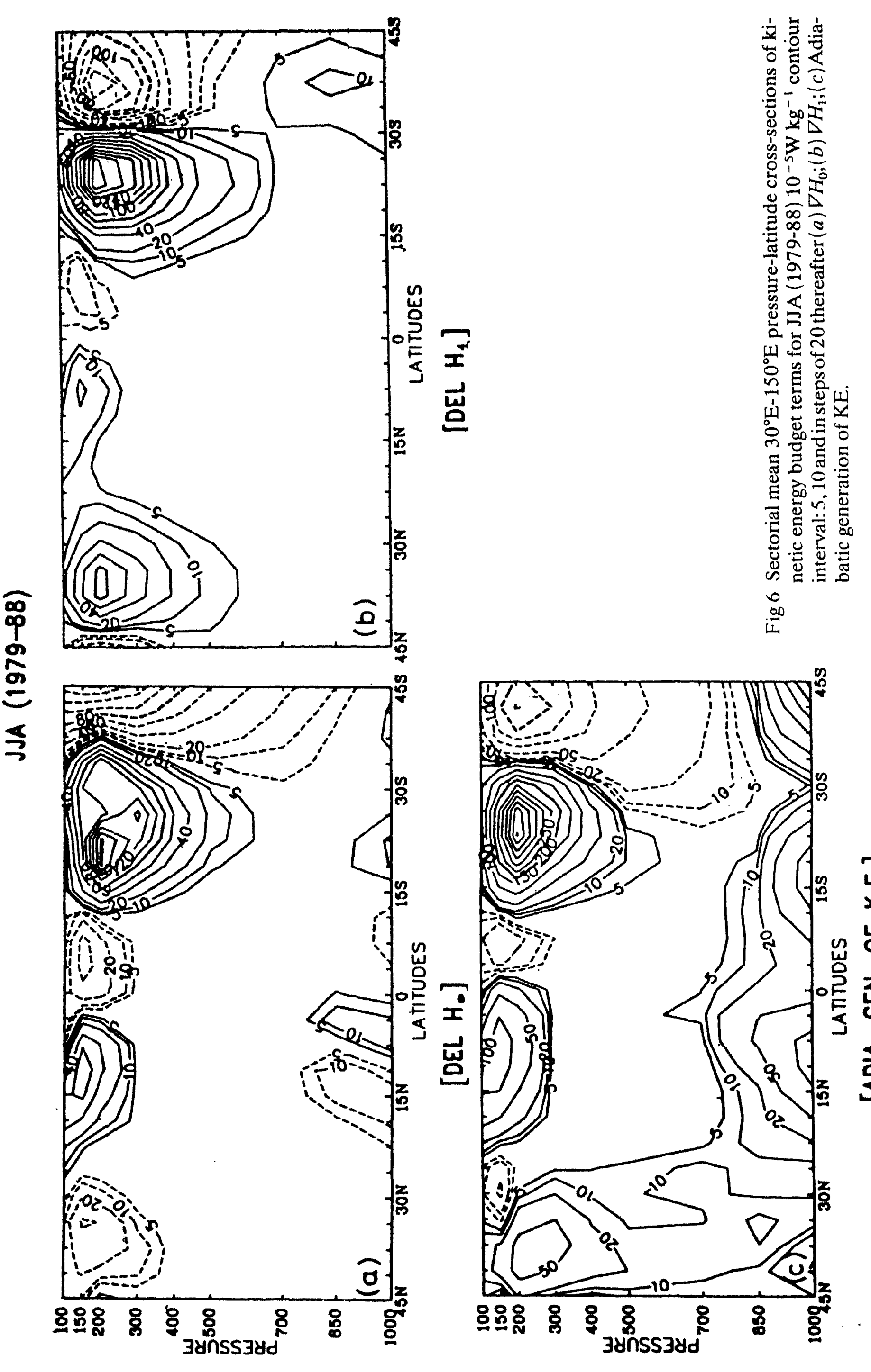


Fig 6 Sectorial mean 30°E-150°E pressure-latitude cross-sections of kinetic energy budget terms for JJA (1979-88) 10^{-5}W kg^{-1} contour interval: 5, 10 and in steps of 20 thereafter (*a*) ∇H_0; (*b*) ∇H_1; (*c*) Adiabatic generation of KE.

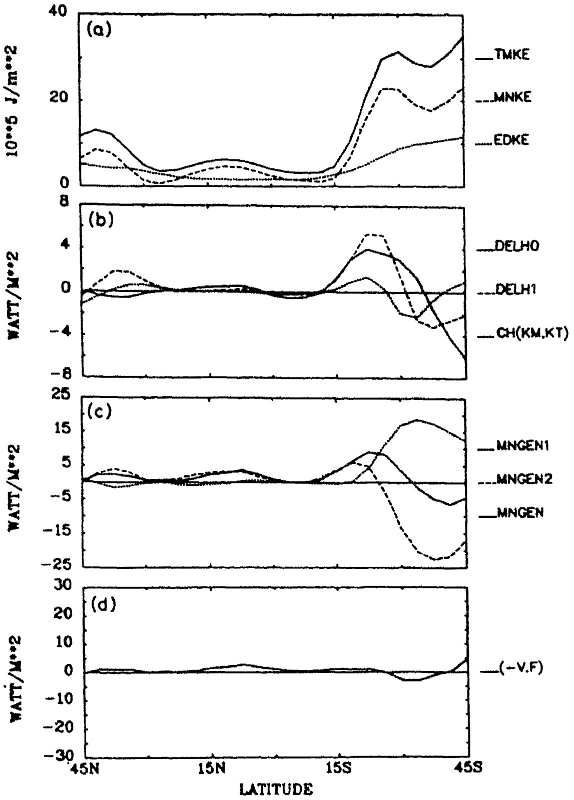


Fig 7 Latitudinal variation of net tropospheric kinetic energy budget terms for JJA (1979-88) (*a*) KE 10^5 Jm^{-2}; (*b*) $\nabla \cdot H$ [Wm^{-2}]; (*c*) Adiabatic generation of KE; (*d*) Dissipation of KE.

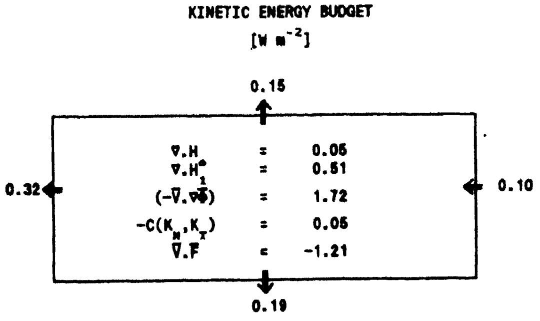


Fig 8 Schematic diagram of the kinetic energy budget for JJA (1979-88) (Wm^{-2})

presence of low level Somali Jet at 850hpa and upper level TEJ at 150hpa. The variation of ΔH_0 over the Northern Hemisphere shows a net kinetic energy flux divergence in the tropics and convergence in the extra tropics. But the variation of ΔH_1 over the Northern Hemisphere indicates a weak flux divergence in the tropics and stronger flux divergence over the extra tropics (Fig. 7*b*). The exchange of kinetic energy between the mean flow and the transients is found to be very small over the NH-tropics. In the Southern Hemisphere, both the flux transport terms and the term representing the exchange of energy between the mean flow and the transients are found to be smaller over the equatorial sector and further south, a strong flux divergence of kinetic energy is observed over the tropics and strong flux convergence over the extra tropics.

Generally, a net adiabatic production of kinetic energy is observed in the tropics and a net destruction over the extra tropics as the tropics are the source regions of kinetic energy due to the presence of direct Hadley circulation cells. Fig. 7*c* demonstrates that the kinetic energy flux transport (import) is compensated with an adiabatic production (destruction) of kinetic energy. Further, it is found that a major contribution to the adiabatic production of kinetic energy in the Northern Hemisphere is accounted by the meridional component of its production. The action of the zonal component of kinetic energy production leads to the net destruction of kinetic energy. However, a clear picture of the reversal in the roles of zonal and meridional components of kinetic energy production is emerged in the Southern Hemisphere to the south of 20°S. Similar results are also found by Kung[3] over the North America. The variation of the net tropospheric kinetic energy dissipation is found to be fairly significant over the tropics (Fig. 7*d*). Only over the sub-tropics of the Southern Hemisphere, negative dissipation phenomena (sub-grid scale processes supplying the energy for maintaining the kinetic energy balance) is observed.

The volume integrated values of the kinetic energy budget over the summer monsoon domain (30°E-150°E; 15°S-30°N) along with the boundary fluxes are shown in Fig. 8. They show that the production of kinetic energy and the dissipation are the most dominant terms. A notable feature found in the analysis is that the flux transport of kinetic energy by the transient eddies dominating the transport by the mean flow itself. A small portion of mean flow kinetic energy is being transferred to the transient eddies. Boundary fluxes signifies that monsoon atmosphere acts as the source of kinetic energy with a considerable outward flux along its western boundary.

Vorticity Budget

In general, among the various vorticity budget terms, vorticity transport and the generation terms (stretching and tipping terms) are the significant terms in the vorticity budget (Holopainen and Oort[37]; Chu *et al.*[14]).

The planetary vorticity advection and the relative vorticity advection terms largely determine the balance of vorticity in the vertically integrated domain (Holopainen and Oort[37]). Hence, the geographical distribution of the vertically integrated vorticity advection terms are shown in Fig. 9. Both of these terms are found to be large in magnitude and opposite in sign over the regions under

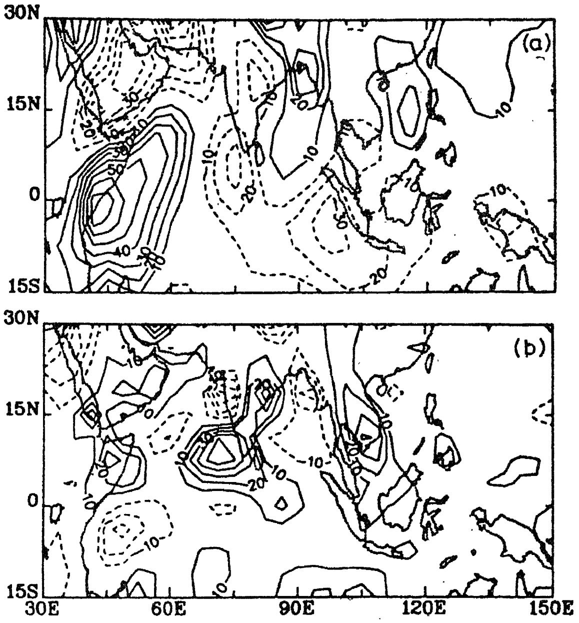


Fig 9 Geographical distribution of vertically integrated vorticity budget terms for JJA (1979-88) (10^{-8} Nm^{-3}) contour interval: 10, (*a*) Planetary vorticity advection; (*b*) Relative vorticity advection.

the influence of the summer monsoon. The planetary vorticity advection (Fig. 9*a*) is found to be dominant over the western Indian Ocean and the adjoining African continent, the regions surrounding the head Bay of Bengal and China Sea. The most prominent zone of planetary vorticity flux divergence over the western Indian Ocean is mainly attributed to the strong cross-equatorial meridional monsoonal flow in the lower troposphere. Over the eastern Arabian Sea, Indian Ocean sectors and the Indian sub-continent, a net convergence of planetary vorticity is observed. All these zones of planetary vorticity convergence over the south peninsula and adjoining seas are characterized by strong relative vorticity divergence (Fig. 9*b*) resulting to a net vorticity advection over the region. In order to maintain the net vorticity advection, vorticity is to be generated within the domain itself for the circulation to sustain.

Sectorial mean vertical cross-section of the absolute vorticity advection term is illustrated in Fig. 10*a*. It depicts that the summer monsoon circulation is characterized by a strong vorticity divergence in the lower troposphere and vorticity convergence in the upper troposphere. While the lower tropospheric vorticity divergence is comparatively weaker over the SH-tropics, the upper tropospheric convergence is found to be stronger in comparison with their corresponding features of the Northern Hemisphere. The lower tropospheric vorticity advection over the NH-tropics is supported by the generation of vorticity by the stretching term due to the presence of cyclonic circulation in association with strong low level convergence (Fig. 10*b*). In the upper levels, the

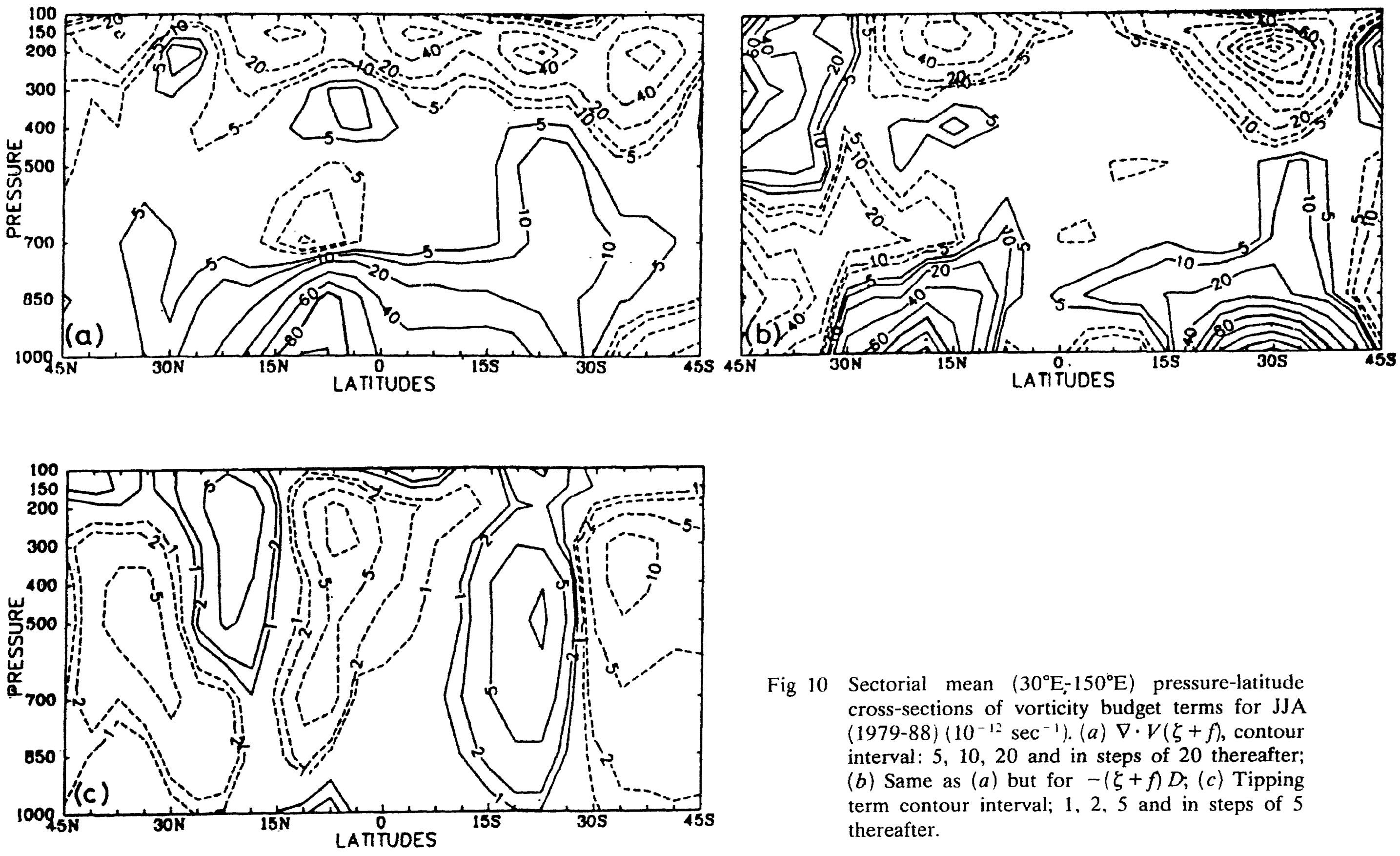


Fig 10 Sectorial mean (30°E-150°E) pressure-latitude cross-sections of vorticity budget terms for JJA (1979-88) (10^{-12} sec^{-1}). (*a*) $\nabla \cdot V(\zeta + f)$, contour interval: 5, 10, 20 and in steps of 20 thereafter; (*b*) Same as (*a*) but for $-(\zeta + f)D$; (*c*) Tipping term contour interval; 1, 2, 5 and in steps of 5 thereafter.

stretching term is contributing towards the generation of anti-cyclonic vorticity to the south of the Tibetan anti-cyclone due to the presence of anti-cyclonic nature of circulation in association with strong divergence over the prevailing easterly wind regime. Over the extra tropics of the Northern Hemisphere, an exactly reverse phenomena are observed i.e., low level anti-cyclonic and upper level cyclonic vorticity generation. To the north of the Tibetan anti-cyclone in the upper troposphere, convergence of the flow is observed in association with the subsidence prevailing in the westerly wind regime. This phenomenon results into the formation of a zone of divergence and anti-cyclonic vorticity production in the lower tropospheric levels overlying the sub-tropical high pressure cell regions. Similar flow characteristics are observed in the sub-tropical high pressure regions to the south of equator. The production of vorticity by the stretching term is found to be minimum over the equatorial latitudes. The solenoidal/tipping term shows an interesting characteristics during the summer monsoon season. Its distribution (Fig. 10*c*) shows a zone of anti-cyclonic vorticity generation over the equatorial latitudes (15°N-7.5°S) flanked by the zones of cyclonic vorticity generation in both the hemispheres. Over the extra tropics, however, the tilting term is contributing to the production of anti-cyclonic vorticity in both the hemispheres.

Latitudinal distribution of net tropospheric vorticity budget terms are depicted in Fig. 11. It is interesting to note that both the planetary vorticity advection and the relative vorticity advection contribute towards the net transport of vorticity over the equatorial sector of the Northern Hemisphere (0°-15°N). However, to north of 15°N and over the entire Southern Hemisphere,

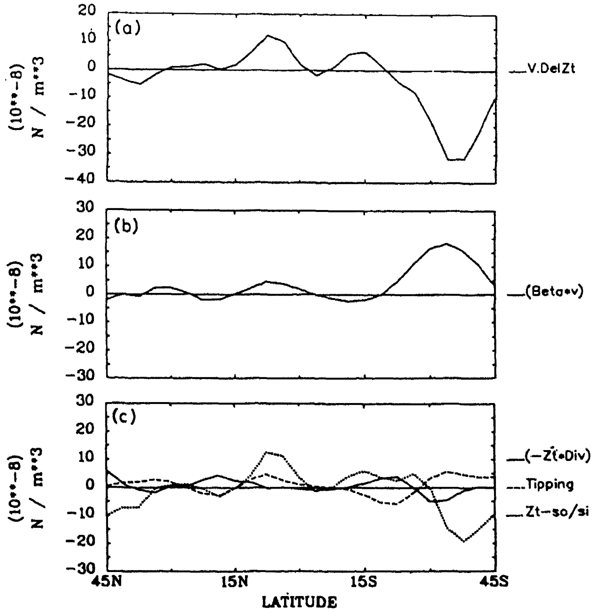


Fig 11 Latitudinal distribution of net tropospheric vorticity budget terms (10^{-8} Nm^{-3}), (*a*) $V\cdot\nabla\zeta$; (*b*) βv; (*c*) Divergence, tipping and residual terms

both the terms tend to strongly oppose each other. The production of vorticity by the stretching and tilting terms is not significant enough to offset the role of vorticity advection terms particularly over the NH-tropics. Over the summer monsoon regime, the balance requirements of the vorticity budget is met through the vertical transport of positive vorticity by the sub-grid scale cumulus convection.

Schematic distribution of volume integrated vorticity budget terms along with the boundary fluxes for the summer monsoon season is shown in Fig. 12. The relative vorticity advection is found to be dominant over the summer monsoon region with a strong transport of vorticity from the domain. The planetary vorticity term also contributes towards the vorticity transport from the domain. The depletion of vorticity by the transport terms is found to be not fully compensated by the vorticity production through the stretching and tilting terms during the summer monsoon season. In our earlier study (Ramesh[38]), it was found that the vorticity production by the stretching and tilting terms was fully compensated only after the summer monsoon gets established over the Indian sub-continent (after middle of July). Till then, the vorticity balance was essentially maintained by the source of positive vorticity from the sub-grid scale processes such as cumulus convection. Many earlier studies (Williams and Grey[39]; Reed and Johnson[40]; Chu *et al.*[14]) have shown that the cumulus convection in the tropics could generate a large apparent source of positive vorticity. Boundary fluxes of vorticity also show maximum inflow from the eastern boundary and maximum outflow through its western boundary during the northern summer.

Angular Momentum Budget

In this section, angular momentum budget analysis is carried out involving the Ω-momentum flux, mean momentum flux divergence, zonal pressure gradient and the mountain torque terms. For the summer monsoon season (JJA), these terms are computed from the 10-year mean fields of the ECMWF analyses.

Vertically integrated horizontal distribution of momentum flux transport and the production terms are presented in Fig. 13. It shows a zone of strong flux divergence over the Arabia extending up to most parts of Indian peninsula

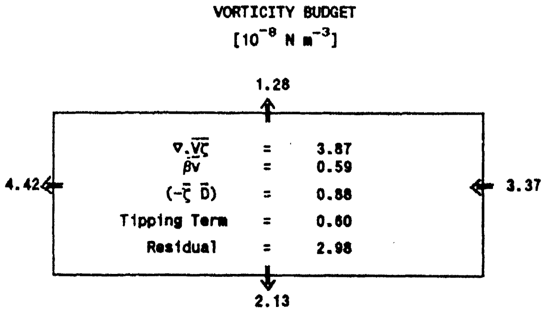


Fig 12 Schematic diagram of the vorticity budget for JJA (1979-88)(10^{-8} Nm^{-3})

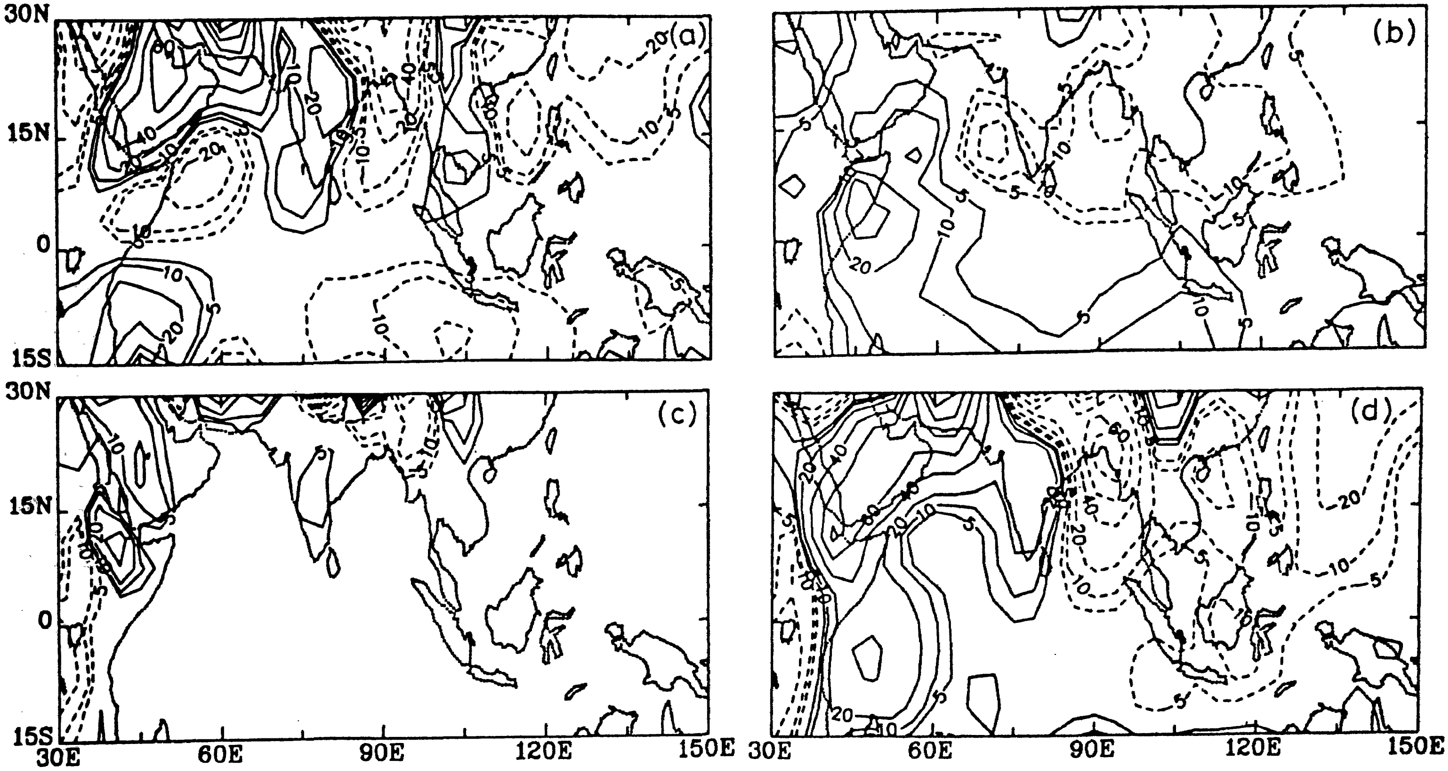


Fig 13 Geographical distribution of vertically integrated angular momentum budget terms for JJA (1979-88) (10^5 kg s^{-2}), contour interval: 5, 10, 20 and in steps of 20 thereafter (*a*) Ω-Momentum flux; (*b*) Relative momentum flux; (*c*) Mountain torque; (*d*) Pressure torque

including eastern Arabian Sea. Over the Tibet and the Gangetic plains and adjoining North East India and Burma and the West Arabian Sea, a zone of Ω-flux convergence is found. Interestingly, the distribution of mean relative momentum flux opposes the Ω-momentum strongly over the Arabian Sea resulting to a widespread zone of momentum flux convergence starting from the western Arabian Sea to the Bay of Bengal across the south peninsula. A net momentum transport is observed over the Arabia and adjoining middle east regions up to Central India. This depletion of momentum is compensated by the production of momentum from the pressure gradient term (Fig. 13*d*). Over the other areas of flux convergence, a strong destruction of momentum by the east-west pressure gradient term is observed. Mountain torque also contributes for the generation of momentum over the peninsular India (Western Ghats and Deccan Plateau), Nepal (Himalayas), Afghanistan (Hindukush ranges) and North African Highlands. However over the hills of North India and North East India, the mountain torque contributes towards the destruction of momentum.

Zonal mean vertical cross-section of the Ω-momentum flux divergence is presented in Fig. 14*a*. The lower tropospheric levels of the northern hemispheric tropics are characterized by a strong flux convergence of momentum and the corresponding levels of the SH-tropics are characterized by a strong flux divergence. A zone of flux divergence is found at middle and upper tropospheric levels of the Northern Hemisphere. But a strong flux convergence exists over the upper troposphere of the Southern Hemisphere. A strong lower tropospheric momentum flux convergence is necessary to maintain the lower tropospheric monsoon westerlies against the loss of momentum through the surface friction during the summer monsoon. Distribution of mean momentum flux transport (Fig. 14*b*) shows an exactly opposite pattern to that of Ω-flux transport in the upper tropospheres of both the hemispheres. However, a prominent zone of momentum flux convergence is observed in the lower troposphere of th NH-tropics (10°N-30°N). Latitudinal height cross-section of mountain torque (Fig. 14*c*) shows a characteristic zone of momentum generation in the lower tropospheric levels (5°N-22.5°N) and destruction over the rest of northern hemispheric latitudes. The zonal east-west pressure gradient distribution over the NH-tropics (Fig. 14*d*) shows a strong destruction of momentum in the lower troposphere and a strong production in the upper troposphere over the NH-tropics.

Latitudinal variation of the mass integrated momentum budget terms are illustrated in Fig. 15. It is evident that the advection terms of the momentum budget contribute to the net inflow of momentum into the summer monsoon regime where low level westerlies are present. This would enable to maintain the strength of the lower tropospheric monsoon westerlies crucial for the transport of moisture against the surface friction. It is also found that both the mountain torque and the pressure gradient terms contribute to the net production of momentum over the summer monsoon region.

Mass integration of the various terms of angular momentum budget is carried out over the monsoon region and the results are presented in Fig. 16. It is found that the earth's momentum flux divergence, the east-west pressure gra-

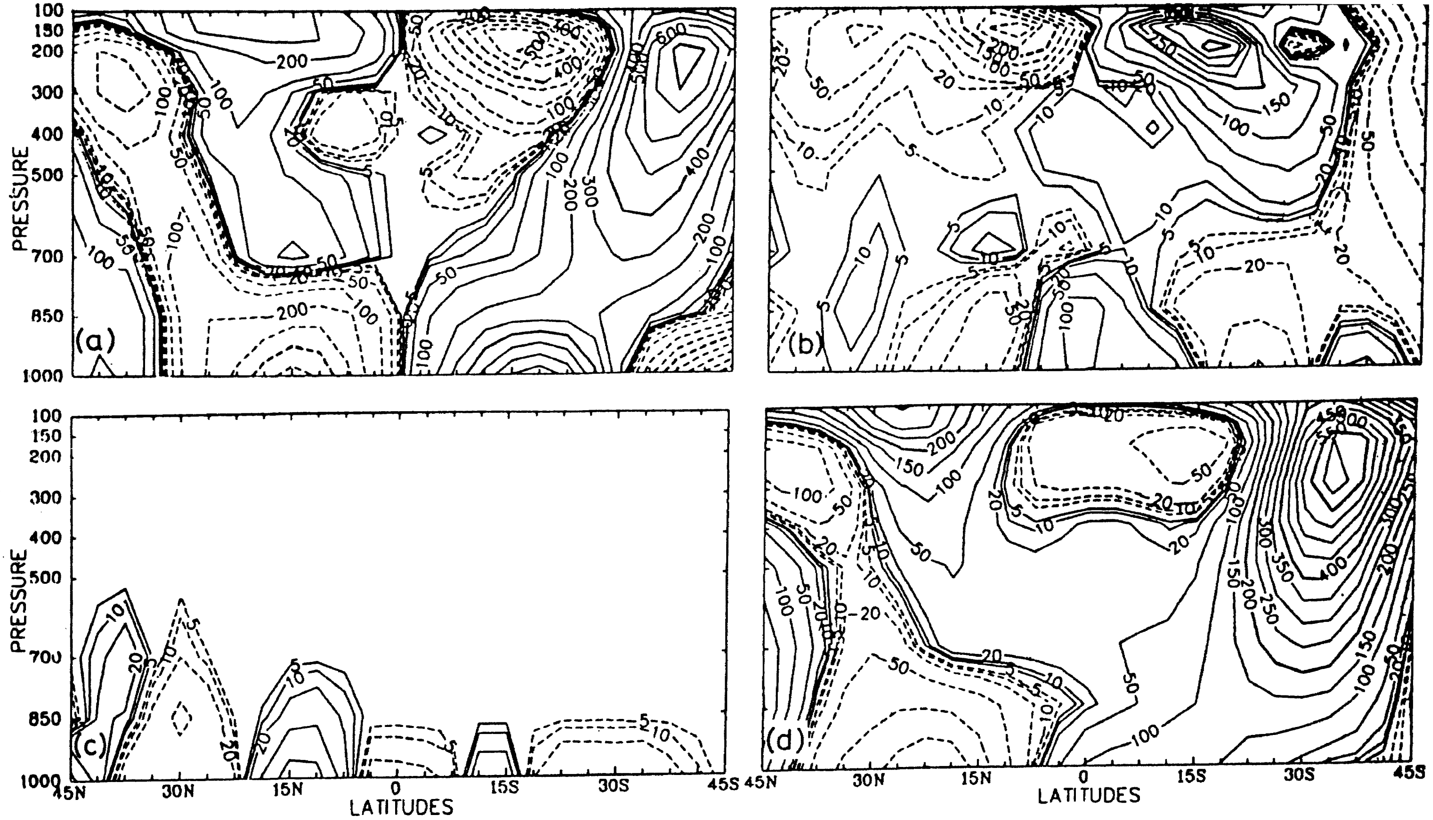


Fig 14 Sectorial mean (30°E-150°E) pressure-latitude cross-sections of angular momentum budget terms for JJA (1979-88) (10^{14} m^4s^{-2}), contour interval: 5, 10, 20, 50 and in steps of 50 thereafter (*a*) Ω-Momentum flux; (*b*) Relative momentum flux; (*c*) Mountain torque; (*d*) Pressure torque

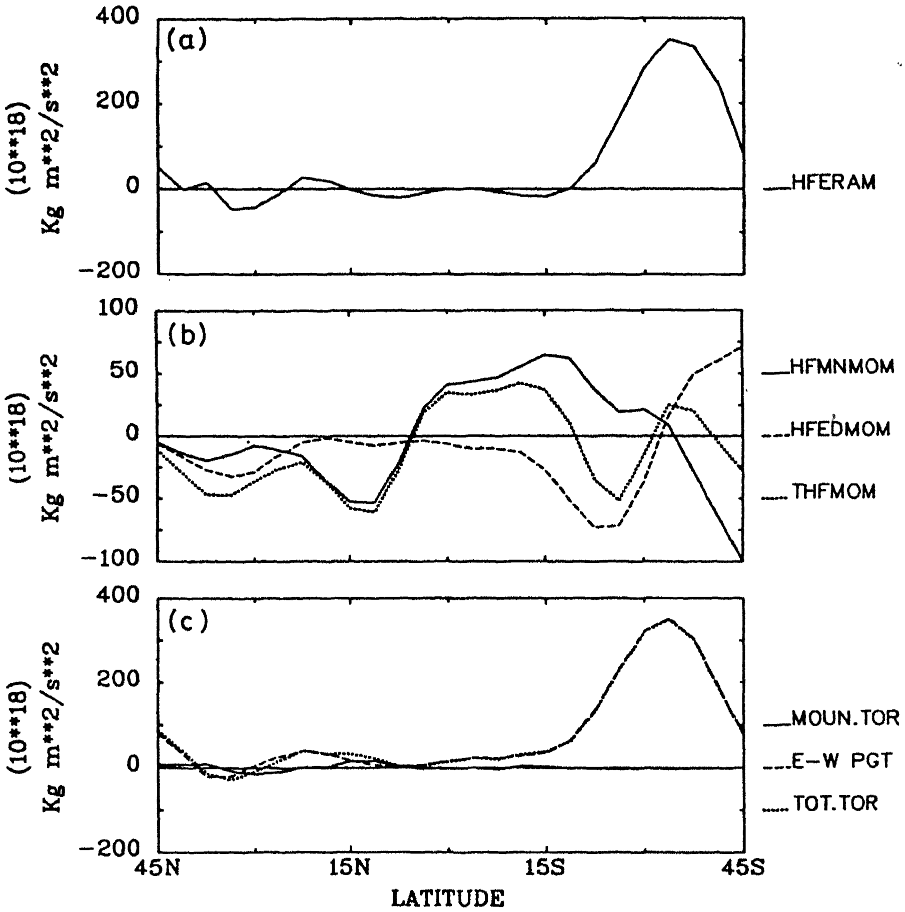


Fig 15 Latitudinal distribution of net tropospheric angular momentum budget terms for JJA (1979-88) (10^{18} kg m^2 s^{-2}) (*a*) Ω-Momentum flux; (*b*) Relative momentum flux; (*c*) Mountain torque, Pressure torque and Frictional torque

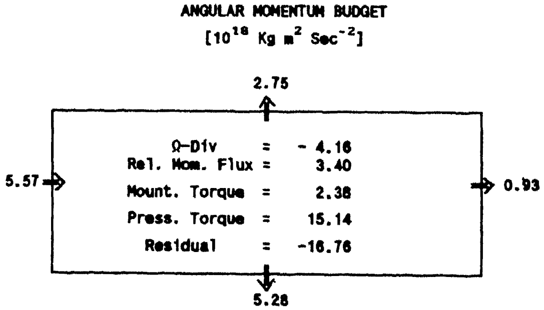
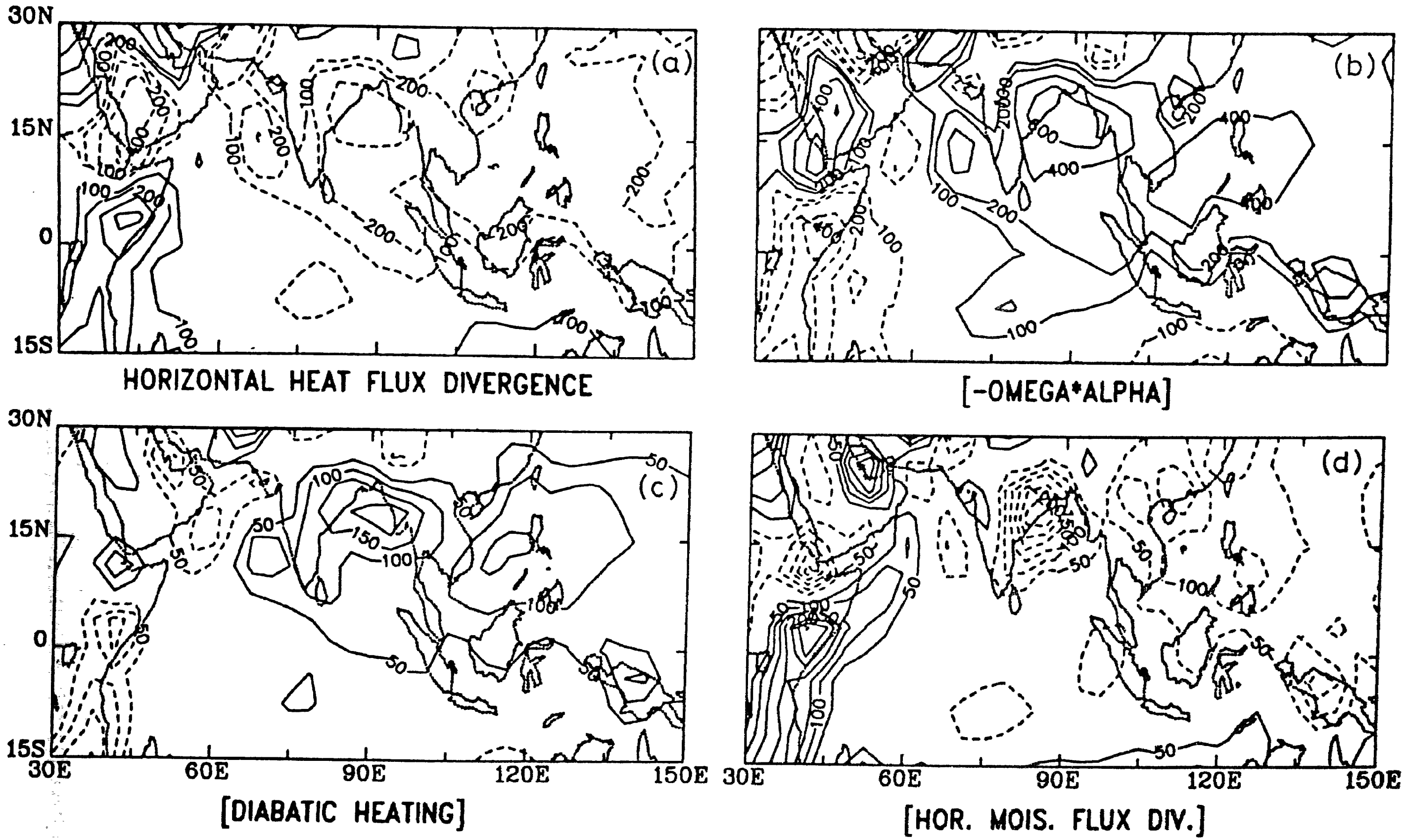


Fig 16 Schematic diagram of angular momentum budget for JJA (1979-88) (10^{18} kg m^2 s^{-2})



HORIZONTAL HEAT FLUX DIVERGENCE

[-OMEGA*ALPHA]

[DIABATIC HEATING]

[HOR. MOIS. FLUX DIV.]

Fig 17 Geographical distribution of vertically integrated heat and moisture budget terms for JJA (1979-88) (W m^{-2}), (*a*) Horizontal heat flux divergence, contour interval: 100, 200 and in steps of 200 thereafter; (*b*) Same as (*a*) but for Adiabatic generation; (*c*) Diabatic beating, contour interval: 50; (*d*) Same as (*c*) but for horizontal moisture flux divergence.

dient and the frictional torque (computed as a residue) are the most dominant terms. Further, it is found that the net effect of momentum transport and the zonal pressure gradient terms is to generate the zonal momentum over the monsoon region which is essential for the maintenance of the lower tropospheric monsoon westerlies against a strong frictional dissipation. During the summer monsoon season, the mountain torque also contributes towards the production of momentum. Boundary fluxes of momentum show maximum inflow from the western boundary and a net outflow of momentum from its southern and northern boundaries.

Heat and Moisture Budgets

Some of the significant features of the summer monsoon circulation are further discussed through the analysis of vertically integrated large scale budgets of heat and moisture over the monsoon region (Fig. 17). Distribution of heat flux divergence (Fig. 17*a*) shows that the summer monsoon region is characterized by flux convergence with two maxima respectively over the Bay of Bengal and the eastern Arabian Sea. A zone of strong heat flux divergence is found over the east African region. The conversion of APE to kinetic energy through adiabatic processes (Fig. 17*b*) exhibits zones of maxima over the Bay of Bengal, eastern Arabian Sea and the western Pacific indicating the presence of strong rising motion over the entire summer monsoon regime. The geographical distribution of the net diabatic heating is also positive all over the region under the influence of the summer monsoon. A primary maximum of heat source is observed over the Bay of Bengal, adjoining Burma and other two heat source regions are found over the eastern Arabian Sea and the West Pacific respectively. Further, all these zones of diabatic heating over the South East Asia are the regions characterized by strong flux convergence of moisture (Fig. 17*d*). In the earlier studies (Mohanty *et al.*[21,22,23]), it was found that the moiture flux convergence precedes the establishment of the diabatic heat source over the monsoon regime. Thus, it may be concluded that the moisture flux convergence plays an important role in determining the diabatic heating patterns crucial for the maintenance of summer monsoon circulation.

The vertical cross-section of the horizontal heat flux divergence (Fig. 18*a*) shows a strong heat flux convergence over the lower and middle tropospheric levels extending right up to 300hpa with a zone of strong heat flux divergence in the upper levels of the NH-tropics (10°N-30°N). But over the Southern Hemisphere, an exactly opposite feature with low level heat flux divergence and upper level flux convergence is observed. No organized distribution is observed over the equatorial latitudes. It was earlier found that the horizontal and vertical heat flux divergence terms are very large in terms of magnitude and to a large extent counter balance each other (Mohanty *et al.*[21]). In the summer monsoon domain, vertical flux divergence is found to dominant in the lower troposphere and horizontal flux divergence is dominant in the upper levels. Adiabatic conversion term shows a zone of energy conversion all through the troposphere with a maximum at around 300hpa over the NH-tropics. Zones of production of APE through a process of kinetic energy conversion are found over the entire Southern Hemisphere (0°-40°S) and the extra tropics

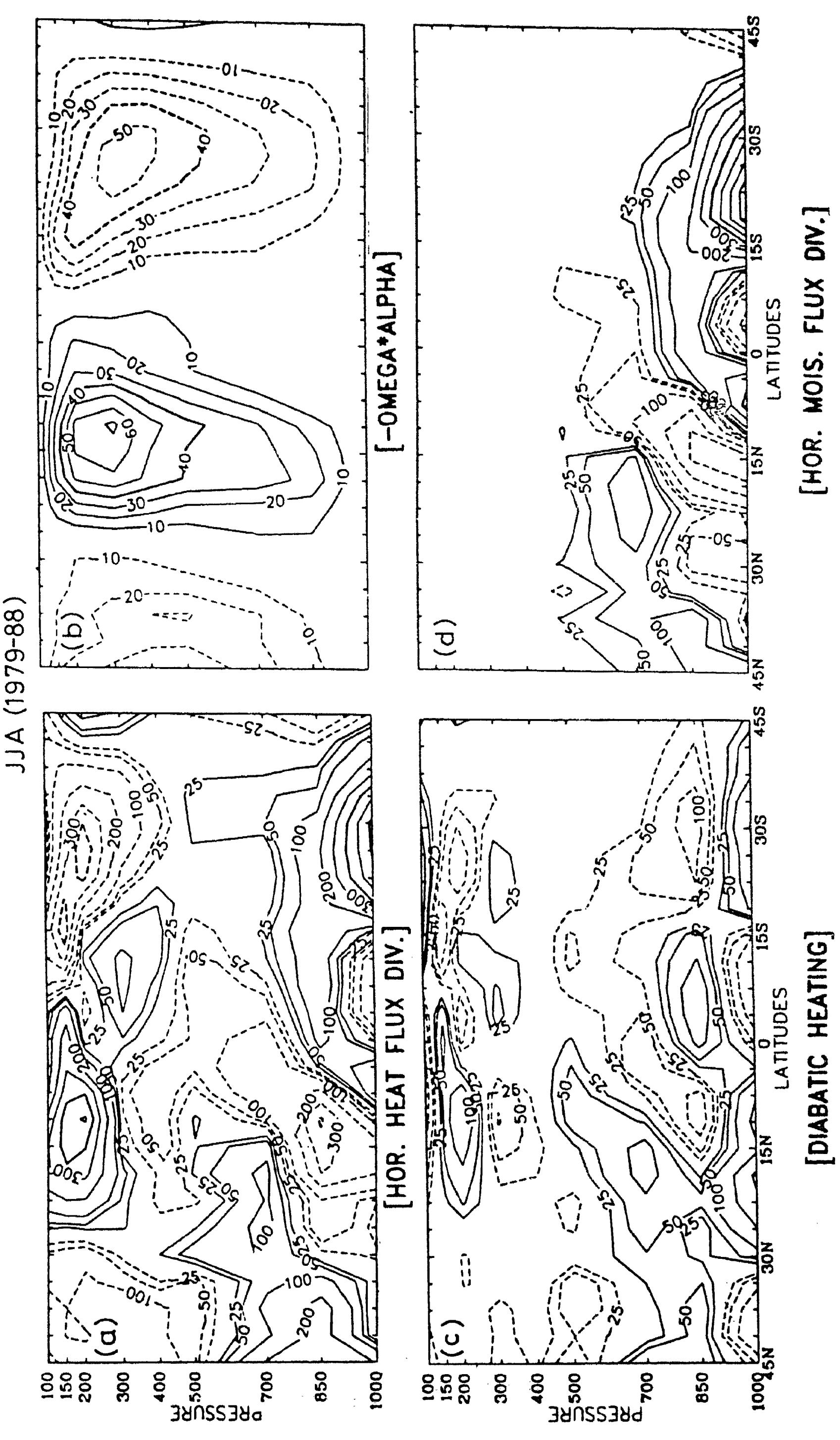


Fig 18 Sectorial mean 30°E-150°E) pressure-latitude cross-sections of heat and moisture terms for JJA (1979-88) (W kg^{-1}) (*a*) Horizontal heat flux divergence, contour interval: 25, 50, 100 and in steps of 100 thereafter; (*b*) Adiabatic generation, contour interval: 10; (*c*) Same as (*a*) but for Diabatic heating; (*d*) Same as (*a*) but for Horizontal moisture flux divergence.

of the Northern Hemisphere. These results reveal that the release of APE at large scales as a source of kinetic energy as a result of rising motion is the characteristic feature of the monsoon circulation. Further, the level of maximum in the upper troposphere agrees well with the earlier finding of energy conversion over the tropics (Nitta[41]; Wallace[42]). Diabatic heating distribution in the sectorially averaged domain (Fig. 18*c*) is characterized by a lower tropospheric maximum close to the ground (1000-850hpa). This maximum of diabatic heating close to the ground may be attributed to the strong heat transfer from the relatively warm ocean and land surfaces by turbulent exchange mechanism. A secondary zone of diabatic heating maximum is also found in the tropical upper troposphere which could be accounted for the strong ascent of moist air and the release of latent heat on a macroscopic scale over the monsoon region.

Horizontal flux transport of moisture (Fig. 18*d*) shows a zone of dominant flux convergence in the lower tropospheric levels of the NH-tropics (10°N-30°N) while a strong flux divergence of moisture is present in the corresponding levels of the Southern Hemisphere during the northern summer season (JJA). These results indicate that the horizontal flux divergence of moisture plays an important role in the monsoon activity by supplying excess moisture to the region which releases diabatic heating in the atmosphere through the process of condensation.

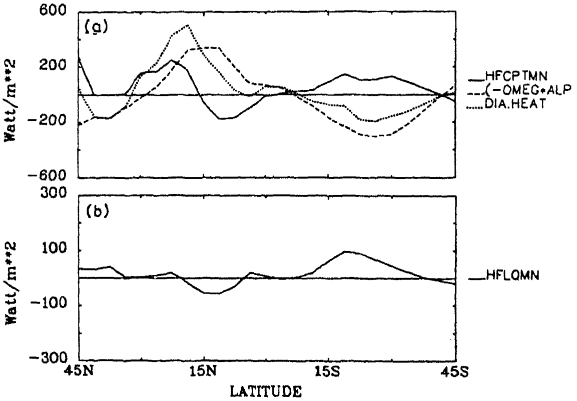


Fig 19 Latitudinal distribution of net tropospheric heat and moisture budget terms for JJA (1979-88) (Wm^{-2}) (*a*) Hor. Heat Flux Div., Adiabatic generation, Diabatic heating: (*b*) Hor. Mois. Flux Div.

Latitudinal distribution of net tropospheric heat and moisture budget terms are depicted in Fig. 19. Summer monsoon atmosphere in the equatorial tropics (0°-18°N) of the Northern Hemisphere is characterized by a net convergence of heat and moisture fluxes into the domain and a net flux transport to the north of it. The adiabatic conversion term also contributes to the production of heat energy over the summer monsoon latitudes leading to the development of a strong diabatic heat source over the region.

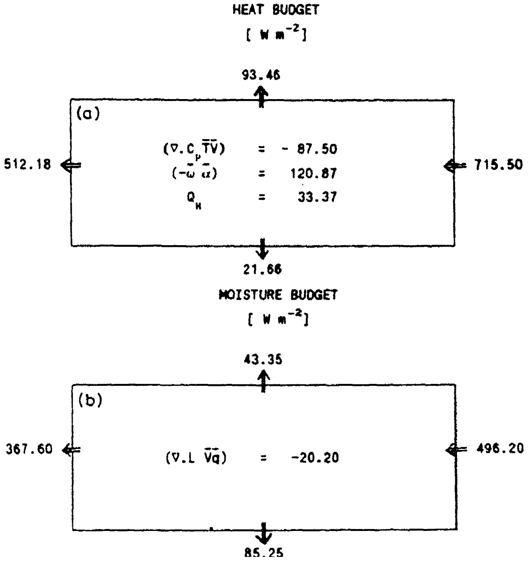


Fig 20 Schematic diagrams of (*a*) Heat Budget and (*b*) Moisture Budget for JJA (1979-88) (Wm^{-2}).

Volume integrated heat and moisture budget analysis for the summer monsoon circulation is presented in Fig. 20. As seen earlier, the summer monsoon is characterized by a net import of sensible heat and latent heat fluxes in to the domain through its eastern boundary. A net outflow of sensible and latent heat is observed through all its other lateral boundaries. Also, a considerable amount of energy is converted from the APE source through the adiabatic processes. The net effect of the transport and adiabatic conversion terms lead to the formation of a strong diabatic heat source in the monsoon atmosphere. The physical mechanism for this heating could be attributed, mainly to the strong turbulent heat transfer from the relatively warm ocean and land surfaces having strong low level winds. Further, considerable latent heat release in the middle and upper troposphere through the intense convective processes during the monsoon season also provides significant contribution towards the net diabatic heating.

Conclusions

Based on the results presented in this study, the following general conclusions may be drawn:

The dominant characteristics of the divergent circulation of the summer monsoon constitute a major center of low level convergence and an upper level divergence over the South East Asia with a large horizontal velocity potential gradients during the summer season.

The geographical region under the influence of the summer monsoon is characterized by tropospheric warming associated with maximum moisture build-up. Such a distribution is favourable for the development of intense convective activity resulting in to copious rainfall over the Indian sub-continent.

The circulation over the East African continent and adjoining western Indian Ocean determines the strength of the cross-equatorial flow and the ageostrophic motions in maintaining the lower tropospheric features of the monsoon. Also, it controls the quantum of heat and moisture flux transport in to the monsoon regime.

The circulation over the Bay of Bengal and the eastern Arabian Sea largely determines the upper tropospheric features of the summer monsoon regime. The extrance/exit regions of the TEJ are characterized by production/destruction of KE which is essential to maintain outflow/inflow prevailing at the respective locations of the TEJ. It is interesting to note that the eastern Arabian Sea maximum of KE production is maintained by the zonal component of the ageostrophic flow while that of the Bay of Bengal is maintained by the meridional component of the ageostrophic flow.

Mean circulation of the summer monsoon is characterized by a net vorticity advection from the region of its influence. This observation indicates that the whole area is quite unstable during the summer monsoon season with the production of vorticity with in the domain itself for maintaining the circulation.

In this analysis, a net momentum flux convergence is observed over the region of low level monsoon westerlies. Such a distribution is essential to maintain the low level monsoon westerlies against the surface friction. It is also found that both the mountain torque and pressure torque contribute towards the production of momentum over the summer monsoon region.

In the summer monsoon domain, a net convergence of heat and moisture is observed. Also, a considerable amount of heat energy is generated through the action of adiabatic processes. The combined effect of these processes lead to the formation of strong diabatic heat sources in this region to maintain the monsoon circulation.

Boundary fluxes of the monsoon regime signifies that the monsoon circulation acts as a source of kinetic energy, vorticity, heat and moisture with a considerable inward (outward) fluxes along its eastern (western) boundaries. In contrast to this as expected from dynamical considerations, boundary fluxes of momentum show maximum inflow from the western boundary and outflow from its southern and northern boundaries.

Acknowledgements

Authors are grateful to the ECMWF, Reading, UK for providing access to their data archives for undertaking this study. We also express our sincere thanks to Professor P K Das, FNA for his valuable suggestions.

References

1 R P Pearce *Quart J R Met Soc* **104** (1978) 737-755
2 E C Kung and P J Smith *Bull Am Met Soc* **55** (1974) No. 7, 768-777
3 E C Kung *Mon Wea Rev* **98** (1970) 911-916
4 E C Kung *Quart J R Met Soc* **97** (1971) 61-74
5 T C Chen *Tellus* **34** (1982) 55-62
6 A H Oort and T C Chan *PAGEOPH* **115** (1977) 1167-1186
7 T S S Anjaneyulu *Quart J R Met Soc* **97** (1971) 103-109
8 R Ananthakrishnan and R N Keshavamurty *Indian J Met & Geophys* **24** (1973) 235-244
9 R N Keshavamurty and S T Awade *Mon Wea Rev* **98** (1970) No. 4, 315-320
10 S S Singh, A H Master and A A Kulkarni *MAUSAM* **35** (1984) 11-16
11 G S Mandal, S C Gupta and S K Saha *MAUSAM* **35** (1984) 145-150
12 J R Holton and D F Colton *J Atmos Sci* **29** (1972) 1124-1128
13 J S Fein *PAGEOPH* **115** (1977) 1493-1500
14 J H Chu, M Yanai and C H Sui *J Met Soc Japan* **59** (1981) No. 4, 535-546
15 S M Daggupaty and D R Sikka *J atmos Sci* **33** (1977) 773-792
16 Y P Rao *Indian J Met & Geophys* **12** (1961) 413-418
17 M Sankar Rao and R Ramanatham *J atmos Sci* **20** (1963) 350-353
18 R N Keshavamurty *Mon Wea Rev* **96** (1968) 23-31
19 C W Newton *J atmos Sci* **28** (1971*a*) 623-628
20 C W Newton *J atmos Sci* **28** (1971*b*) 1329-1341
21 U C Mohanty, S K Dube and P C Sinha *MAUSAM* **33** (1982*a*) 139-152
22 U C Mohanty, S K Dube and P C Sinha *MAUSAM* **33** (1982*b*) 285-294
23 U C Mohanty, S K Dube and M P Singh *J Met Soc Japan* **61** (1983) 208-221
24 R P Pearce and U C Mohanty *J atmos Sci* **41** (1984) 1620-1639
25 A C Lorenc *Mon Wea Rev* **109** (1981) 701-721
26 L Bengtsson, M Kanamitsu, P Kallberg and S Uppala *Bull Am Met Soc* **63** (1982) 29-43
27 P Lonnberg and D Shaw *Met Bull M*1.5/1 *Res Manl* **1** (1983)
28 M Kanamitsu *Dynamic Meteorology—Data Assimilation Methods Springer Verlag* (1981)
29 B J Hoskins, H H Hsu, I N James, M Masutani, P D Sardeshmukh and G H White *WCRP*-27 *WMO/TD No.* 326 (1989) pp 1-162
30 P R Julian *GWE Newslett No.* 3 *USC-GARP* (1984) 20-22
31 O'Brien *J appl Met* **9** (1970) 197-203
32 U C Mohanty, K J Ramesh and S K Dash *J Met Soc Japan* **67** (1989) No. 5, 691-704
33 R E Newell, J W Kidson, D G Vincent and G J Boer *The General Circulation of the Tropical Atmosphere and Interactions with Extratropical Latitudes* Vol. 1 The MIT Press (1972)
34 Y P Rao *South-West Monsoon, Met Monograph No* 1/1976 IMD India
35 P K Das *Monsoons WMO Monograph* (*Fifth IMO Lectures*) 1984
36 T N Krishnamurti *J atmos Sci* **28** (1971) 1342-1347
37 E O Holopainen and A H Oort *J atmos Sci* **38** (1981) 262-269
38 K J Ramesh *A Study on Certain Aspects of the Dynamics and Energetics of the Asian Summer Monsoon* (Ph.D. Thesis) IIT Delhi (1990)
39 K T Williams and W M Grey *Tellus* **25** (1973) 313-336
40 R J Reed and R H Johnson *J atmos Sci* **31** (1974) 1784-1790
41 T Nitta *J Met Soc Japan* **48** (1970) 524-528
42 J M Wallace *Rev Geophys Space Phys* **9** (1972) 557-612

CHAOS AND PREDICTABILITY IN FORECASTING THE MONSOONS

T N PALMER

ECMWF Shinfield Park, Berkshire RG2 9AX, Reading, England

(*Received* 5 *June* 1993; *Accepted* 17 *June* 1993)

Day-to-day evolution of weather is only predictable up to ten days or so in advance. This can be understood in terms of the chaotic nature of atmospheric dynamics. On the other hand, skilful longer range predictions of seasonally averaged rainfall have been successfully made, as exemplified by the monsoon forecasts from the India Meteorological Department. This dichotomy is explored using the prototype Lorenz 3-component chaotic model. The model is applied to the Indian monsoon problem, with break and active monsoon periods corresponding to the Lorenz-model regimes. The dichotomy is resolved by taking into account the impact of lower boundary forcing anomalies on the regime probability density function. By studying the impact of such anomalies in a chaotic model, it can be seen that seasonal prediction is necessarily probabilistic, and requires ensemble integration techniques when tackled using dynamically-based forecast models.

Key Words: **Chaos; Predictability; Forecast; Monsoon Breaks; Lorenz 3-Component Chaotic Model; Butterfly Effect; Tropical Convergence Zone (TCZ)**

Introduction

The unpredictability of weather is legendary, and despite the best efforts of man and technology, reliable predictions of the day-to-day evolution of weather can only be made up to a week or so ahead.

The scientific reasons for such unpredictability are now well established. Despite well formulated and precisely known deterministic laws of motion, the atmosphere is a chaotic system, whose evolution is sensitive to starting conditions. As a result, the influence of the smallest imaginable error in the initial conditions for a weather forecast will eventually have an influence on the evolution of daily weather. This sensitivity is vividly characterized by the "butterfly effect"; the flap of a butterfly's wings in Delhi, could make the difference between a fine or stormy day in London a month or so later.

The first precise mathematical model that demonstrated the existence of chaos in a simple nonlinear dissipative system was formulated by Lorenz[1]. His three component equations

$$\dot{X} = -\sigma X + \sigma Y,$$

$$\dot{Y} = -XZ + rX - Y \qquad \ldots (1)$$

and

$$\dot{Z} = XY - bZ$$

are not a rigorous truncation of the equations for the large-scale atmosphere, nevertheless the behaviour of these equations is a good qualitative guide to the predictability of weather.

Superficially, it would appear that beyond the limit of predictability of day-to-day evolution of weather, there is no prospect of skilful atmospheric prediction. Despite this, longer-range predictions are routinely made. Most famous of these are the seasonal monsoon predictions of the India Meteorological Department now made for over 100 years using models based on empirical relationships between monsoon and worldwide climatic predictors. Whilst there is no doubt that these predictions are skilful[2] the question arises as to why such long-range predictions are possible in a fundamentally chaotic atmosphere.

To resolve this conflict, the dynamics of the Lorenz model are discussed in detail in the next section, particularly the differences between predictions of the first and second kind. Then in the subsequent two sections an attempt is made to use the model to study, conceptually at least, the predictability of the Indian monsoon both on timescales of up to 10 days and timescales of a season. Some practical remarks on the development of dynamical models for long-range monsoon prediction are made in the last section.

Properties of the Lorenz Model

Fig. 1 shows a number of integrations of the Lorenz model (eq. 1) in the phase space spanned by the three model variables. The lighter background shows the famous Lorenz attractor: the model climate. This attractor is clearly inhomogeneous; it has two distinct regimes. Coincidentally, the attractor itself has a butterfly shape.

In Fig. 1 are shown the result of three ensemble integrations of the Lorenz equations. The ensembles start from different parts of the attractor. The ensemble of initial conditions, a small ring of points, can be thought of as representing some uncertainty in the initial state. The three panels show that the unpredictability asociated with finite-time integrations of the Lorenz model depends very strongly on the initial conditions themselves. For example, in the first panel all members of the ensemble of points make a transition from the left hand regime to the right hand regime. In this sense the regime transition is very predictable.

In the second and third panels the evolution of the Lorenz state vector is much less predictable. In the second case one could say that there was about a 40% chance of a regime transition, whilst in the last case there is virtually nothing that can be said about the likely final state.

It is fairly easy to understand qualitatively the behaviour of these ensemble integrations in the Lorenz model. In particular it can be seen that most of the ensemble dispersion occurs at the middle bottom of the attractor, in a neighbourhood of the origin of phase space ($X=Y=Z=0$). This in turn is associated with a particularly unstable fixed point of the equations at the origin. (There are also unstable fixed points near the regime centroids; however, compared with the origin, these instabilities are not particularly severe).

Following the terminology of Lorenz[3], we can refer to these types of initial value forecasts as "predictions of the first kind". By contrast, "predictions of the

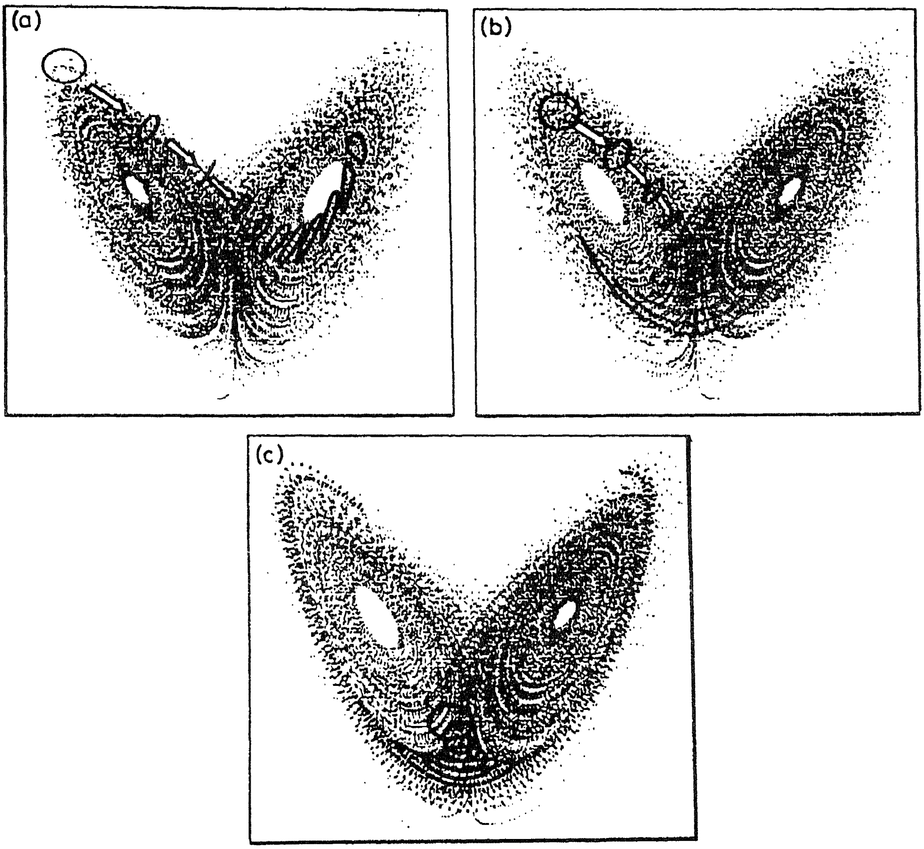


Fig 1 Phase-space evolution of an ensemble of initial points on the Lorenz attractor, for three sets of initial conditions, superimposed on the Lorenz attractor

second kind" are not initial value problems, rather they are concerned with how the attractor might change as some system parameter is increased. (For example, how climate changes as atmospheric CO_2 is doubled is a prediction of the second kind).

Understanding the difference between predictions of the first and second kind are important for a conceptual appreciation of the difference between monsoon prediction on the daily and on the seasonal timescale (*see* the next sections); we therefore discuss these issues further.

We can study a prediction of the second kind by introducing some fixed normalised "forcing" into the Lorenz equations

$$\dot{X} = -\sigma X + \sigma Y + \alpha F_X$$

$$\dot{Y} = -XZ + rZ - Y + \alpha F_Y$$

and

$$\dot{Z} = XY - bZ + \alpha F_Z \qquad \ldots (2)$$

and ask how the Lorenz climate changes as α is increased from zero.

The result of one such experiment is given in Fig. 2 in terms of the probability density function (PDF) associated with the state vector of the Lorenz

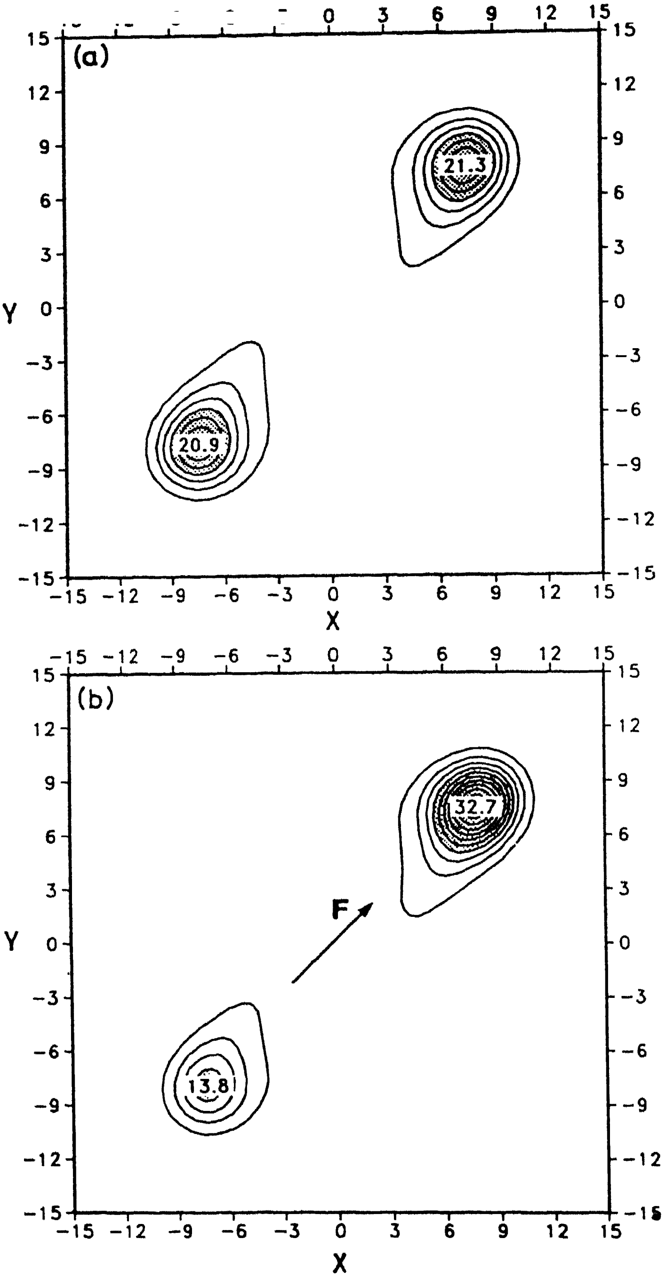


Fig 2 Probability density function of the Lorenz model, low-pass filtered to remove oscillations around a regime centroid. *a*) from the unforced model. *b*) with a constant forcing in the X-Y plane pointing between the regimes.

model. In the original model (1), the PDF is symmetric between regimes, so that the probability of finding the system in one regime is equal to the probability of finding it in the other regime.

The influence of a forcing vector which points from one regime to the other, on the state vector PDF, is shown in Fig. 2*b*. It can be seen that the probability of finding the system in the positive *X*-*Y* regime is now larger than that of finding it in the other regime.

This result may seem straightforward, but there are two subtleties. Firstly note that the effect of forcing is not to translate the whole attractor in the direction of forcing. The phase-space coordinates of the regime centroids are largely unchanged by the forcing. The reason for this is that the effect of the forcing is felt most keenly by the state vector when it is in a neighbourhood of the unstable fixed point near the origin (just as errors in initial conditions amplify most strongly in the vicinity of the origin). When the state vector resides in a regime, the effect of the forcing is relatively small. (This is analogous to the effect of applying a small force to a pencil balanced on its tip—the effect of the small force can be quite large. Applying the same small force when the pencil is falling has much less effect).

The second important point to note about this numerical experiment is that in both Figs 2*a* & *b*, the evolution of the state vector is chaotic. In other words predictions of the first kind in either (1) or (2) are limited by sensitivity to errors in the initial state. However, despite this sensitivity to initial state, the effect of the applied forcing in (2) is itself predictable in terms of changes to the PDF of the system.

Monsoon Regimes

What have these remarks to do with atmospheric prediction, and with monsoon forecasting in particular? As mentioned in the introduction, the Lorenz equations are not themselves truncations of the equations for the large-scale atmosphere. On the other hand they do have qualitative similarities. For example, the inhomogeneous regime structure of the real climate is well known in studies of the atmosphere. For example, in higher latitudes (e.g., over London), the weather can persist in an unsettled state for many weeks. During this period, the winds are predominantly from the west, generally strong, and carry with them rain-bearing weather fronts. Suddenly the character of the weather will change corresponding to a settled spell with light or even easterly winds. In winter this settled weather can be rather cold; in summer it is generally warm. These qualitative remarks can be made precise using cluster analysis techniques (e.g., Mo and Ghil, 1988).

Is there similar regime-like behaviour during the summer monsoon? It is well known[5] that within the summer monsoon season, the large-scale rainfall oscillates between active spells with good rainfall, and weak spells with little rainfall. These weak spells are often referred to as "monsoon breaks". In general such large-scale rainfall is associated with the so-called "tropical convergence zone" (TCZ), a region where lower tropospheric winds are convergent. For Indian longitudes, there are two favourable locations of the TCZ; one associated with convergence over the heated continent, giving the active monsoon phase, the other associated with convergence over the equatorial Indian ocean, giving the break phase. As shown by Sikka and Gadgil[6], the PDF of the TCZ is bimodal. The evolution between active and break periods is aperiodic; but with a preferred timescale of a few weeks. Typically the transition time between break and active spells is shorter than the residence time of the spells themselves.

Whilst comprehensive general circulation models of the atmosphere appear able to simulate such intraseasonal variability[6], large-scale numerical weather prediction models (such as the European Centre for Medium Range Weather Forecasts-ECMWF-model) have difficulty in predicting transitions, e.g. from active to break phase, beyond the first week of prediction.

Based on these similarities, and others presented below, the present author would like to put forward the Lorenz model paradigm in discussing long-range monsoon predictability. In this paradigm, the two regimes of the Lorenz model correspond to the active and break periods of the monsoon. To be definite, let us suppose that the positive X-Y regime in Fig 2*a* corresponds to the oceanic TCZ, with reduced monsoon rainfall, and that the negative X-Y regime corresponds to the continental TCZ with enhanced monsoon rainfall. A simplification of the weather forecast problem applied to the monsoon would be the prediction of regime transitions, from active to break period or *vice versa*. (Of course in the first day or so of the forecast period, one is able to predict much more detail; perhaps this simplified view is more appropriate between about 5 and 10 ahead). The seasonal forecasting problem, on the other hand, is not concerned about predicting specific regime transitions, but rather about forecasting changes to the probabiltiy of occurrence of the strong and weak monsoon regimes.

Lower Boundary Forcing

We now introduce a further important concept. Atmospheric motion is bounded below by the surface of either land or ocean (or indeed ice). Heat and momentum are "exchanged" through this surface. The temperature of the surface is therefore an important boundary condition that influences atmospheric circulations. Sea surface temperature (SST) is a particularly important boundary condition in that it is coherent over large spatial scales, and it generally varies slowly on the timescale of individual weather events, but is not constant from year to year. The particular importance of the lower boundary conditions on monsoon predictability has been discussed by Charney and Shukla[7].

Since, as far as the atmosphere itself is concerned, SST is an external variable, the influence of changing SST on the monsoon can in some ways be thought of as a prediction of the second kind. Now it is known from observations that year to year variations in tropical Pacific SST associated with the El Nino/Southern Oscillation event have a strong influence on interannual variations in the monsoon[8,9]. Therefore, for conceptual purposes, let the external forcing F in the Lorenz model (2), represent an index of anomalous tropical Pacific SST. (As Webster and Yang[10] have emphasised, in reality the interaction between the monsoon and the El Nino is 2 way).

The paradigm then goes as follows. In a "normal" situation ($F=0$ as in Fig 2*a*) the seasonal mean monsoon rainfall comprises aperiodic active and break spells with a climatological PDF. With El Nino SST (F as in Fig 2*b*), the monsoon continues to fluctuate aperiodically between active and break spells, however, the PDF is now biased towards the break spells, and the time-mean rainfall is lower than normal. During a cold El Nino event (these are sometimes referred to as La Nina events) in which tropical Pacific SST falls by up to several degrees (F reversed from its direction in Fig. 2*b*), the regime PDF is now biased towards the opposite active spell, and time-mean rainfall is higher than normal.

Now these results refer to shifts in the PDF's of a chaotic system. Hence, in any one given season, the best that one can say is that the probability of lower than average rainfall is increased during an El Nino event, or increased during a La Nina event. Because of the chaotic dynamics, there is not a one-to-one

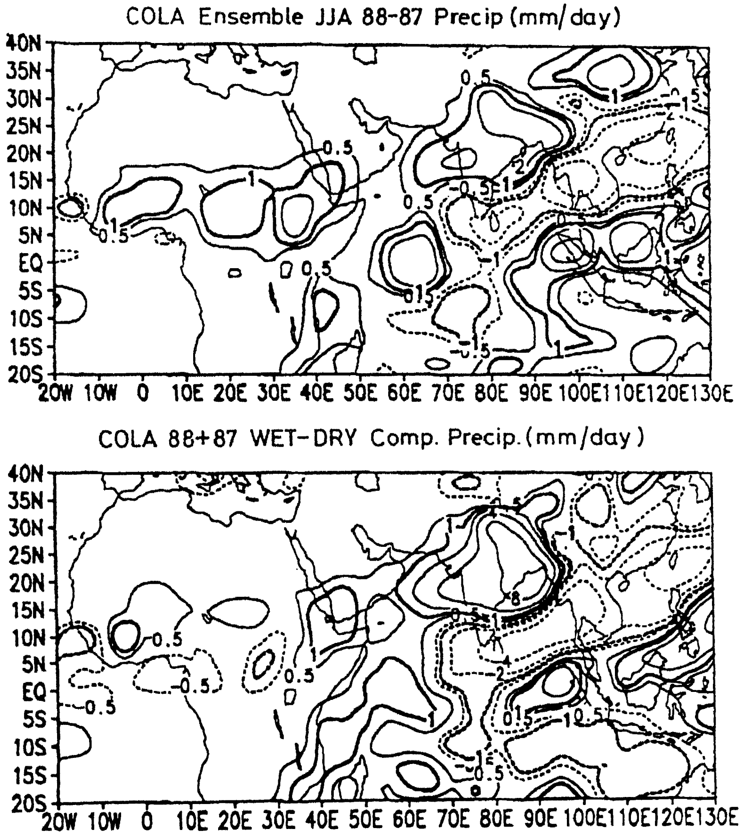

COLA Ensemble JJA 88-87 Precip (mm/day)
40N
35N
30N
25N
20N
15N
10N
5N
EQ
5S
10S
15S
20S
20W 10W 0 10E 20E 30E 40E 50E 60E 70E 80E 90E 100E 110E 120E 130E
COLA 88+87 WET-DRY Comp. Precip.(mm/day)
40N
35N
30N
25N
20N
15N
10N
5N
EQ
5S
10S
15S
20S
20W 10W 0 10E 20E 30E 40E 50E 60E 70E 80E 90E 100E 110E 120E 130E

Fig 3 *a*) Ensemble-mean JJA difference rainfall between seasonal integrations of the University of Maryland COLA general circulation model integrated with observed SST from 1988 and 1987 respectively. The initial conditions for the ensemble were taken from observed analyses from 1, 2 and 3 May of 1988 and 1987 (respectively). *b*) Composite difference between active and break periods of the monsoon from within the seasonal GCM integrations (courtesy of M Fennessy) link between the SST anomalies and monsoon rainfall, and there is still a chance of above average rainfall during an El Nino event.

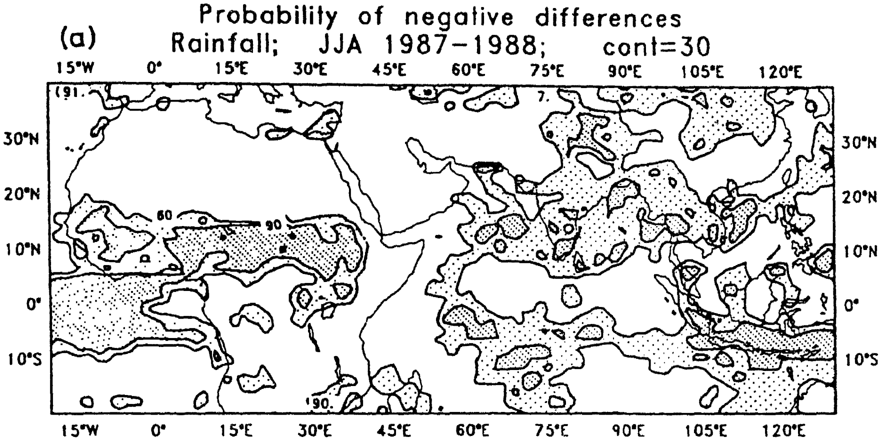
One of the principal testable hypotheses of such a paradigm is that the patterns of interannual fluctuations in monsoon rainfall, should correspond to patterns associated with the active and break spells. Observations[11] and modelling studies (M. Fennessy, *personal commun*) appear to lend some support to this notion. Results from the modelling study of M Fennessy are quite dramatic in this regard, and are shown in Fig. 3. The University of Maryland COLA general circulation model (GCM) was integrated over two summer seasons with observed SSTs for 1987 and 1988. The difference in seasonal mean rainfall (Fig 3a) was compared with the composite difference in rainfall between active and break periods within a season (Fig. 3*b*). Over much of the Indian Ocean and adjoining land masses, the patterns of interannual and intraseasonal rainfall agree closely.

The main purpose of this paradigm is to suggest that whilst the detailed evolution of the monsoon may be nonlinear and chaotic, seasonal prediction may be possible through the influence of lower boundary forcing on the regime PDFs. Again, for emphasis, one important consequence of this is that seasonal predictability will never be perfect.

To illustrate this, consider the following. Suppose we integrate a comprehensive atmospheric weather prediction model from 1 May of some year for four months, to the end of August, and let us also suppose for the sake of argument that we know exactly what the SST anomalies are through the season. Over the simulated summer season June-August (JJA) the model will make a number of transitions between active and break spells, and the JJA seasonal average rainfall is determined by the relative frequency of active to break spells. This, however, is a poorly sampled estimate of the attractor's PDF. Hence, if we now made a second integration from one day later (2 May), the relative frequency of active and break spells in the seasonal integration will not be the same as from the first integration, because of the chaotic nature of the regime transitions. We may have to perform many integrations from different starting conditions, before we can determine reliably the PDF of the active and break regimes, given the SST anomalies. However, any one of the estimated frequencies could correspond to what actually happens.

Such a set of ensemble calculations has been performed using the ECMWF model. The model has been integrated from 9 different atmospheric initial conditions for two contrasting monsoon seasons 1987 and 1988. For each season, and for each pair or ensemble members, the difference between the JJA rainfall in 1987 and 1988 has been calculated. Finally the fraction of positive and negative rainfall differences within the ensembles have been calculated. These are shown in Fig 4.

In Fig 4, it can be seen that for 1987 over India the ensemble indicates a probability of a relatively dry monsoon season generally exceeding 60%. Over the African Sahel, on the other hand, the probability of a relatively weak monsoon in 1987 is generally much higher, in excess of 90%. The stronger probabilities for the Sahel region correspond to the fact that chaotic intraseasonal variability is much smaller in that area than over India.


(a)
Probability of negative differences
Rainfall; JJA 1987-1988; cont=30
15°W
0°
15°E
30°E
45°E
60°E
75°E
90°E
105°E
120°E
30°N
20°N
10°N
0°
10°S
60
90

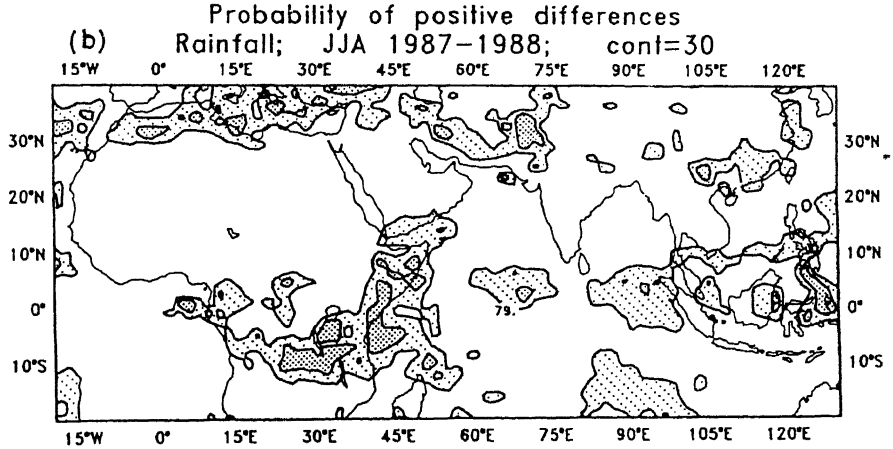

(b)
Probability of positive differences
Rainfall; JJA 1987-1988; cont=30
15°W
0°
15°E
30°E
45°E
60°E
75°E
90°E
105°E
120°E
30°N
20°N
10°N
0°
10°S
79.

Fig 4 Probability estimates of rainfall over parts of the tropics based on 9-member ensembles of 120-day integrations with observed SST, based on the ECMWF numerical weather prediction model. JJA 1987-JJA 1988. Probability (%) that rainfall difference is *a*) negative, *b*) positive. Contours are 60% and 90%.

We have highlighted the role of the tropical Pacific SST anomalies on the monsoon. However, they are not the only important lower boundary anomalies that influence the monsoon. SST's of other ocean basins play a role, as do land surface forcing. It has been known for many years that Eurasian snow cover in the seasons immediately preceding the monsoon can be important, and modelling experiments have confirmed this[12].

Seasonal Monsoon Prediction using Dynamical Models

It has been suggested that despite the chaotic day-to-day evolution of weather, seasonal prediction is possible because of the influence of the lower boundary conditions on the regime PDFs associated with the climate attractor. The relative influence of these boundary conditions is larger in the tropics than in the extratropics for reasons discussed by Charney and Shukla[7]. In the case of the monsoons, these regimes corresponds to active and break periods. In order to be able to assess reliably the impact of lower boundary conditions on the regime PDFs, it is necessary to integrate a weather prediction model many times from different initial conditions.

In some cases, it may be sufficient to persist the lower boundary anomalies occurring at the beginning of the forecast, throughout the whole forecast period. However, in general, it is necessary to be able to predict the evolution of the atmosphere's lower boundary conditions throughout the seasonal forecast period. To do this one must couple the atmosphere to the underlying land surface and ocean, and treat the SST as a truly dynamical internal variable of the system, governed, like the atmosphere, by precisely known physically-based laws of motion[13]. Of course, relative to the whole system, the forecast problem is a prediction of the first kind, though in this case it is the initial values of the ocean (and land surface) that are important (and that determine the limit of predictability), rather than the initial conditions of the atmosphere. Such coupled models are at present at the stage of development, and operational seasonal prediction using such coupled models has not yet begun in earnest. It may take many years before such dynamical models are more skilful than the empirical model predictions of the India Meteorological Office. However, seasonal dynamical monsoon prediction is not many years away, and, chaos not withstanding, our expectations of success are high.

Acknowledgements

The author is thankful to Drs F Molleni and M Fennessy for supplying figs 2 & 3 respectively. Thanks are also due to Shri J Shukla for helpful comments.

References

1 E N Lorenz *J atmos Sci* **42** (1963) 433-471

2 V Gowariker, V Thapliyal, S M Kulshrestha, G S Mandal, N Sen Roy and D R Sikka *Mausam* **42** (1991) 125-130

3 E N Lorenz *Climate Predictability: The Physical Basis of Climate Modelling W M O GARP Publ Sr* **16** (1975) 132-136

4 K C Mo and M Ghil *J geophys Res* **93D** (1988) 10927-10952

5 S Gadgil and J Srinivasan *Met atoms Phys* **44** (1990) 119-312

6 D R Sikka and S Gadgil *Mon Wea Rev* **108** (1980) 1840-1853

7 J G Charney and J Shukla *Predictability of Monsoons: Monsoon Dynamics* (Ed J Lighthilla and R Pearce) Cambridge Univ Press (1981) 735 pp

8 J Shukla and D Paolino *Mon Wea Rev* **111** (1983) 1830-1837

9 T N Palmer, C Brankovic, P Viterbo and M J Miller *J Climatol* 5 (1992) 399-417

10 P J Webster and S Yang *Quart J R met Soc* **118** (1992) 877-926

11 S Gadgil *Aus met Mag* **36** (1988) 193-204

12 T P Barnett, L Dumenil, U Schlese, E Roecker and M Latif *J atoms Sci* **46** (1980) 661-685

13 D L T Anderson and J Willebrand *Rep Prog Phys* **55** (1992) 1-36

14 B N Goswami *Proc Indian natn Sci Acad* **60A** (1994) No 1 101-120

15 T N Palmer *Bull Am met Soc* **74** (1993) 49-65

16 WMO *Simulation of Interannual and Intraseasonal Monsoon Variability* WCRP 68 (1992)

NUMERICAL EXPERIMENTS OF THE INFLUENCES OF ANOMALY HEATING OVER THE TIBETAN PLATEAU ON THE FORMATION AND EVOLUTION OF THE ASIAN SUMMER MONSOON

FAN YUN

Beijing Institute of Meteorology, Beijing-100 081, *PRC*

CHEN LONGXUN

Chinese Academy of Meteorological Science, Beijing 100 081, *PRC*

A three-level global atmospheric general circulation model (AGCM) has been used to study the impacts of the anomaly heating over the Tibetan Plateau on the formation and evolution of the Asian summer monsoon. The control experiment reproduces the most features of the observation in the Asian summer monsoon region. Two numerical experiments are performed with the anomaly snow cover over the Plateau in winter and without atmospheric diabatic heating over the Plateau in summer, respectively. The results show that a larger snow albedo than the normal on the Plateau in winter will keep the ground colder and cause the circulation anomaly. Its effects can expend to the whole summer and make the reduction of precipitation in the middle and lower reaches of Changjiang (Yangtze) River while it enhances in the South China Sea. When we remove the atmospheric diabatic heating over the Plateau in summer, the Indian summer monsoon will disappear and the east Asian summer monsoon will be weakened Therefore, it is revealed that the responses of the summer monsoon circulation and its evolution to the abnormal heating over the Plateau are remarkable.

Key Words: Numerical Experiments; Anomaly Heating; Tibetan Plateau; Asian Summer Monsoon

Introduction

Tibetan (Qinghai-Xizang) Plateau is the largest and highest plateau with the most complex geographical features in the world. The position of the Plateau locates over the zone of westerlies in winter and over the transitional zone of easterlies and westerlies in summer. The earth-atmosphere system near the Plateau is a cold heat source in winter and a heat source in summer, the differences of the circulation patterns over the Plateau and its neighbouring area features of synoptic processes in Asia may be attributed to its existance. Thus the Tibetan Plateau has special dynamic and thermodynamic effects on the global and regional atmospheric motions in different seasons.

There have been a lot of studies on the Tibetan Plateau and its influences on the atmospheric circulation[1-4]. Especially during May to August of FGGE (1979), the Qinghai-Xizang Plateau Meteorological Experiments (QXPMEX) was made by China and a lot of meteorological data were obtained. By using these data, Chinese meteorologists have done a lot of work and our knowledge about the Plateau meteorology has been improved obviously. However, there are many problems still with us.

In this paper, we use a three-level global atmospheric general circulation model (AGCM)[5] to study the anomaly heating effects of the Tibetan Plateau on the formation and evolution of the Asian summer monsoon. Based on this model, three experiments have been conducted. Comparing the results obtained from the experiments, the effects of the heat forcing of the Plateau are discussed.

Brief Discription of the Model and Experimental Schemes

The model we use is a three-level global atmospheric general circulation model including a planetary boundary layer (PBL). The thickness of the PBL is 50 hPa and from the top of PBL to 200 hPa is free atmosphere. The free atmosphere is divided into two levels. For the PBL and the free atmosphere level we use the independent σ-coordinate system respectively (σ' for the PBL and σ for the free atmosphere level):

$$\sigma = \frac{P - P_T}{}$$

and

$$\sigma' = \frac{P - P_s + 50}{50}$$

where P_T is the constant pressure at top of the model ($P_T = 200$ hPa), P_s is the pressure at the earth's surface and P^* is the thickness of the free atmosphere

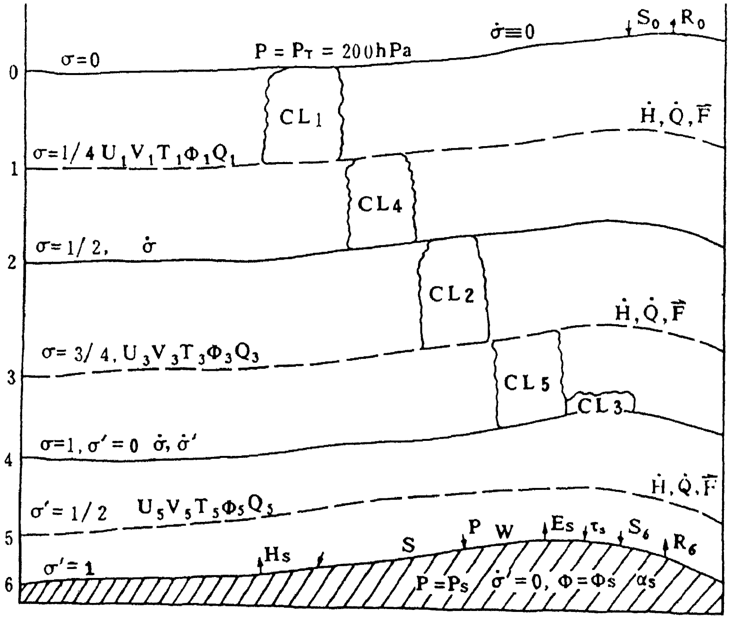

$\dot{\sigma} \equiv 0$
S_0 R_0
$\sigma = 0$
$P = P_T = 200\text{hPa}$
0
CL1
$\dot{H}, \dot{Q}, \vec{F}$
$\sigma = 1/4\ U_1 V_1 T_1 \Phi_1 Q_1$
1
CL4
$\sigma = 1/2,\ \dot{\sigma}$
2
CL2
$\dot{H}, \dot{Q}, \vec{F}$
$\sigma = 3/4,\ U_3 V_3 T_3 \Phi_3 Q_3$
3
CL5
CL3
$\sigma = 1, \sigma' = 0\ \dot{\sigma}, \dot{\sigma}'$
4
$\dot{H}, \dot{Q}, \vec{F}$
$\sigma' = 1/2\ \ U_5 V_5 T_5 \Phi_5 Q_5$
5
P W Es τ_s S_6 R_6
S
Hs
$\sigma' = 1$
6
$P = P_s$ $\dot{\sigma}' = 0,\ \Phi = \Phi_s$ α_s

Fig 1 The vertical structure, principal variables and 5 cloud types of the three-level AGCM

level ($P^* = P_s - P_T - 50$), 50 hPa is the thickness of the PBL. The vertical structure, the basic variables and cloud types of the model are shown in Fig. 1, where u, v, T, q, Φ are the principal variables, P_T the pressure at the model top, P_s the pressure at the earth's surface, S and R the net downward solar and net upward terrestrial radiation at the top of the model atmosphere (subscript 0) or at the earth's surface (subscript 6), as the surface albedo, H, Q and $\mathbf{F}$ the diabatic heating, moisture source and friction, H_s the surface sensible heat flux, P the precipitation rate, E_s the surface moisture flux, τ_s the surface momentum flux, S the snow mass, and w the ground wetness. CL1-CL5 are 5 basic cloud types, 0-4 is free atmosphere level and 4-6 is PBL.

Differential Equations

For free atmosphere we use the σ-coordinates and the horizontal momentum equation in vector form and thermodynamic energy equation can be written in the form:

$$\frac{\partial}{\partial t}(P^*\mathbf{V}) + \nabla(\mathrm{P^*}\mathbf{V}\mathbf{V}) + \frac{\partial}{\partial \sigma}(P^*\mathbf{V}\sigma) + f\mathbf{K} \times P^*\mathbf{V} + P^*\nabla\Phi + P^*\sigma\alpha\nabla P^* = P^*\mathbf{F} \quad \ldots (1)$$

and

$$\frac{\partial}{\partial t}(P^*C_pT) + \nabla\cdot(P^*C_pT\mathbf{V}) + \left(\frac{P}{P_{00}}\right) \times \frac{\partial}{\partial \sigma}(P^*C_p\sigma\upsilon) - P^*\sigma\alpha\left(\frac{\partial P^*}{\partial t} + \mathbf{V}\cdot\nabla P^*\right)$$

$$= P^*H, \qquad \ldots (2)$$

where

$$\upsilon = Y\left(\frac{P_{00}}{P}\right)^x, \quad x = R/C_p = 0.286.$$

The moisture continuity equation is

$$\frac{\partial}{\partial t}(P^*q) + \nabla\cdot(P^*q\mathbf{V}) + \frac{\partial}{\partial \sigma}(P^*q\dot{\sigma}) = P^*\dot{Q} \qquad \ldots (3)$$

The state equation is

$$\alpha = \frac{RT}{P} \qquad \ldots (4)$$

The hydrostatic balance equation is

$$\frac{\partial}{\partial\sigma}(\Phi)+P^*\alpha=0 \qquad \ldots(5)$$

and the mass continuity equation is

$$\frac{\partial}{\partial t}(P^*)+\nabla\cdot(P^*\mathbf{V})+\frac{\partial}{\partial\sigma}(P^*/\dot{\sigma})=0 \qquad \ldots(6)$$

For the PBL we use σ' coordinate in the constant thickness level, the horizontal momentum equation can be written as follows:

$$\frac{\partial}{\partial t}(\mathbf{V})+\nabla(\mathbf{V}\mathbf{V})+\frac{\partial}{\partial\sigma'}(\mathbf{V}\dot{\sigma}')+f\mathbf{K}\times\mathbf{V}+\nabla\Phi+\sigma'\,\alpha\nabla P^*=\mathbf{F}. \qquad \ldots(7)$$

The thermodynamic energy equation is

$$\frac{\partial}{\partial t}(C_pT)+\nabla\cdot(C_pT\mathbf{V})+\left(\frac{P}{P_{00}}\right)^x\frac{\partial}{\partial\sigma'}(C_p\dot{\sigma}'v)-\sigma'\,\alpha\left(\frac{\partial P^*}{\partial t}+\mathbf{V}\cdot\nabla P^*\right)=\dot{H} \qquad \ldots(8)$$

and the moisture continuity equations is

$$\frac{\partial}{\partial t}(q)+\nabla\cdot(q\mathbf{V})+\frac{\partial}{\partial\sigma'}(q\dot{\sigma}')=\dot{Q} \qquad \ldots(9)$$

and the mass continuity equation is

$$\nabla\cdot\mathbf{V}+\frac{}{\partial\sigma'}(\sigma')=0 \qquad \ldots(10)$$

where

$$\nabla(P^*\mathbf{V}\mathbf{V})=(P^*\mathbf{V}\cdot\nabla)\mathbf{V}+\mathbf{V}\nabla+\mathbf{V}\nabla\cdot(P^*\mathbf{V})$$

$$\nabla(\mathbf{V}\mathbf{V})=(\mathbf{V}\cdot\nabla)\mathbf{V}+\mathbf{V}\nabla\cdot\mathbf{V}$$

with $P_{00}=1000$ hPa a reference pressure.

The forcing terms are:

$\mathbf{F}$ = horizontal frictional force vector per unit mass

$\dot{H}$ = diabatic heating rate per unit mass

$\dot{Q}$ = rate of moisture addition per unit mass

The primary dependent variables are:

Φ = geopotential hight

α = specific volume
P = pressure

$\dot{\sigma}=\dfrac{\mathrm{d}\sigma}{\mathrm{d}t}$, $\dot{\sigma}_2=\dfrac{25}{P^*}\nabla\cdot\mathbf{V}_5+\dfrac{1}{4P^*}\nabla\cdot|P^*(\mathbf{V}_3-\mathbf{V}_1)$ Sigma vertical velocity

$\dot{\sigma}_4=\dfrac{50}{P^*}\nabla\cdot\mathbf{V}_5$, $\dot{\sigma}'_4=\nabla\cdot\mathbf{V}_5$

$\mathbf{V}=(u, v)$, horizontal velocity vector

T = temperature

q = water vapour mixing ratio (kg/kg)

Physical Processes

The physical processes in the model include the conditions at the earth's surface, turbulent fluxes in the boundary layer, friction, diabatic heating and moisture sources, cloud types and radiations. The conditions at the earth's surface involve the orography, surface type, surface albedo, ground temperature, ground hydrology, snow mass budget, snowmelt, monthly climatological sea-surface temperatures and sea ice. The turbulent fluxes are momentum, sensible heat and moisture fluxes. The diabatic heating consists of convective adjustment, large-scale condensation and evaporation cumulus convective. For the calculation of the radiation, the concentration of absorption gases in atmosphere, such as water vapour, carbon dioxide and ozone, are all constant, while the amount of high-, medium- and low-level clouds is variable.

The experimental schemes are as the following: (1) Control experiment (Exp. 1). The AGCM has been run for ten model years. We take the last fields from Nov. 1 to next Jul. 31 as the results of the control experiment. (2) Experiment on the anomaly snow cover over the Tibetan Plateau (Exp. 2). In this experiment, we run the AGCM from Nov. 1 to next summer but take the abnormal snow surface albedo over the Plateau between November and December. (3) Experiment without atmospheric diabatic heating over the Tibetan Plateau (Exp. 3). In this experiment, we run the AGCM from Apr. 1 to Jul. 31 but the atmospheric diabatic heating over the Plateau has been removed.

Analysis and Discussion of the Experimental Results

There are a low of experimental results obtained from the tests. In order to save space, the attention is focused on the differences of the air flow at upper and lower level, surface air temperature and precipitation in the Asian summer monsoon region so that we can examine the anomaly heating effects of the Tibetan Plateau on the Asian summer monsoon.

Results of the Control Experiment (*Ex.* 1)

In order to show the ability of the AGCM to reproduce the climate mean

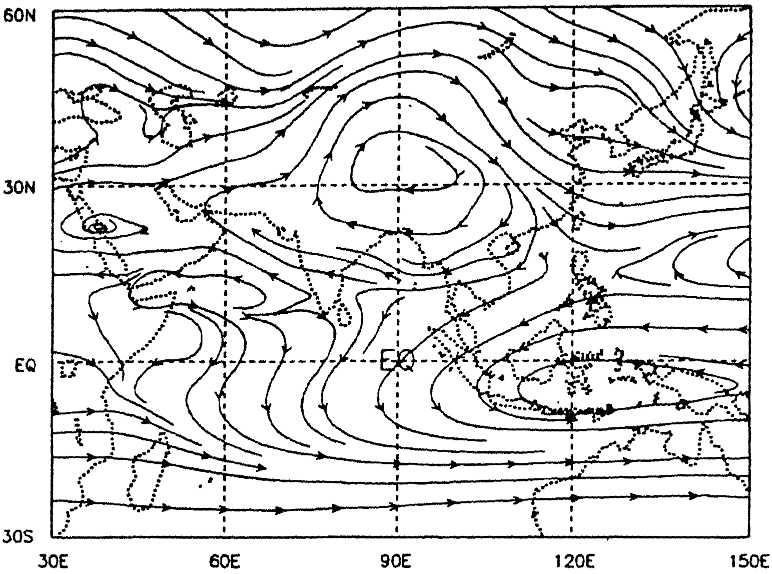


Fig 2(a) The simulated 250 hPa streamline in July

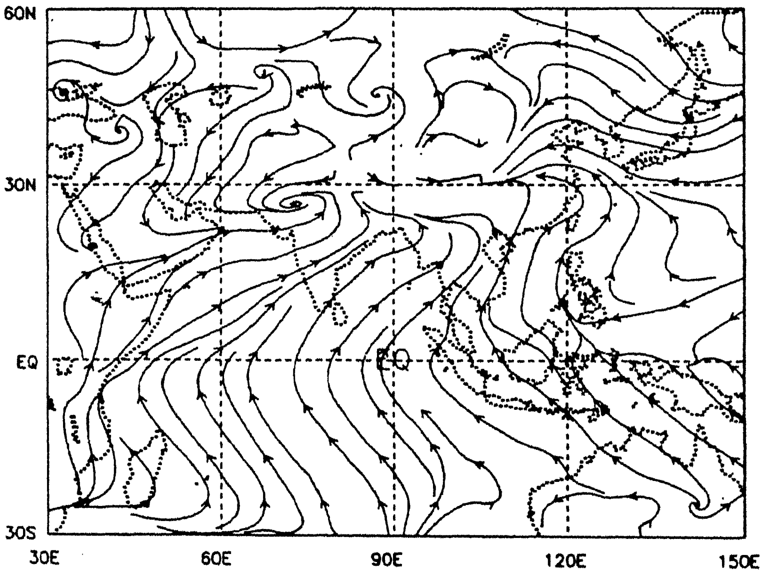


Fig 2(b) The simulated 975 hPa streamline in July

patterns of the Asian summer monsoon, we give some results from the control experiment.

The average July simulated flow patterns are shown in Figs 2*a* and 2*b*. At the upper level (250 hPa), the simulated southern Asian anticyclone is closer to that observed, its center is at 90°C, 32°N where the observed center is at

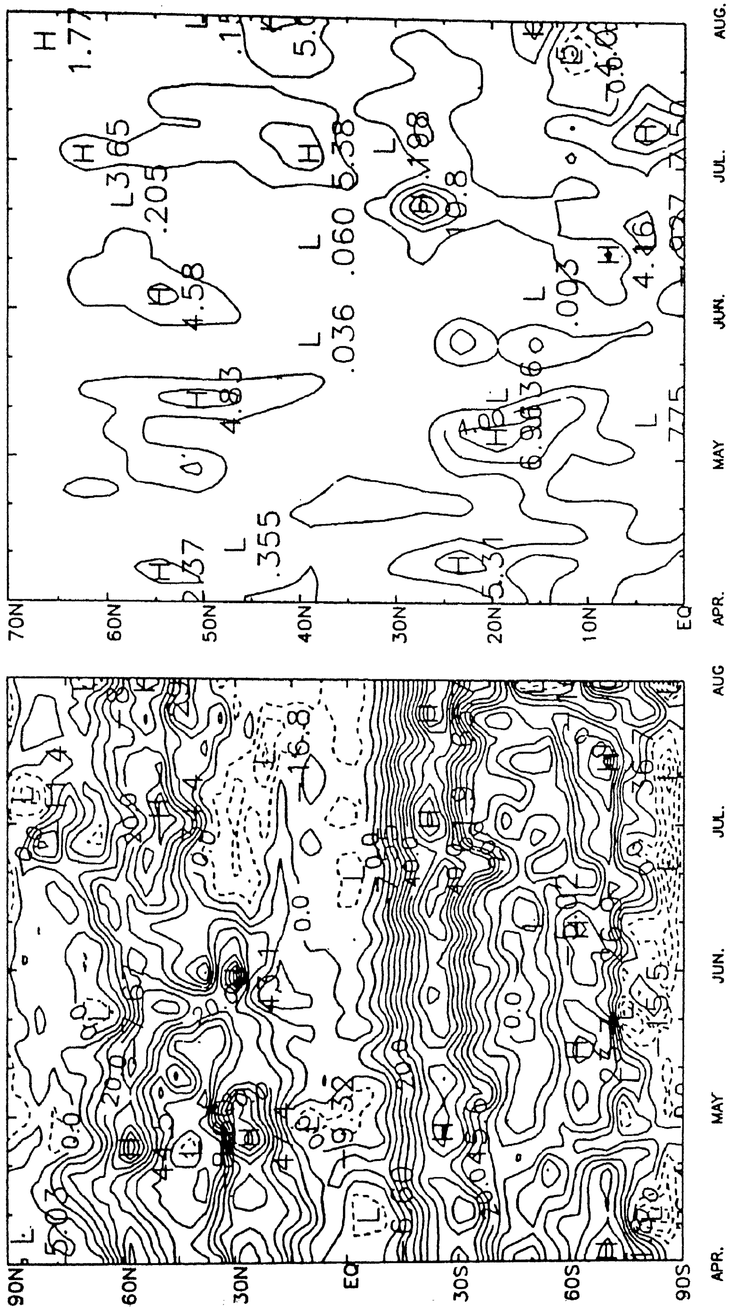


Fig 3 The simulated latitude-time cross-section of the (*a*) zonal wind (at 250 hPa) along 85-90°E and (*b*) precipitation along 110-120°E

88°E, 28°N. The simulated two easterly jets over the south side of Qinghai-Xizang Plateau, the northern one is linked with Asian upper anticyclone and the southern one is joined with the equatorial anticyclone and it is the most intense upper cross equatorial flow, are in better agreement with observation. At the lower level (975 hPa), the simulated low layer Somali jet, monsoon trough, the cross equatorial flows over the South China Sea and Bay of Bengal are in closer agreement with observation. The low pressure belt between Yangtze River and Huaihe River, the cold anticyclone over Australia are successfully simulated. The major features of simulated July surface air temperature, sea-level pressure and precipitation are in closer agreement with the observations.

It is an important symbol for a climate model to simulate the seasonal variation of the flow and precipitation successfully. The lititude-time cross section of the zonal wind (at 250 hPa) along 85-90°E is shown in Figs 3*a* and the latitude-time cross section of rainfal along 110-120°E is shown in Fig. 3*b*. From Figs 3*a*, we see that there are two westerly jets in northern and southern side of Tibetan plateau between April and May, one is the polar-front jet and the other is subtropical jet. Compared with its two sides, the zonal wind speed is weaker over the plateau. In early May, the easterly flow begins across the equator. In early June, the simulated subtropical westerly jet disappears suddenly and the easterly flow appears there immediately. In mid-summer there are two centres in the easterly flow. The result shows that this model can successfully simulate the seasonal variation of the flow at the upper level.

The simulated seasonal variation of precipitation from South China Sea to east of China is shown in Figs 3*b*. There are two rainfall belts, one is the ITCZ in South China Sea and the other is in the mainland. For the former, its centre is near the equator in May and moves northward to 15°N after middle July. This variation is in better agreement with the satellite observation. For the latter, it forms due to the interaction of the summer monsoon and cold air, and has the features of the subtropical rainfall belt. The observation shows that this rainfall belt is in South China during April to June and is called the rain belt in the former flood season (South China Maiyu), it jumps to the middle and lower Yangtze River region in middle June and is called Yangtze River Maiyu, and it again jumps to North China during July and August. From the Figs 3*b* we see that the simulated mainland rainfall belt is in South China (the centre near 22°N) during April to June and is in closer to the actual South China Maiyu, it jumps to near 30°N from middle June to July and is similar to the Yangtze River Maiyu, then it jumps again to 40°N in middle July and is in closer to North China rainfall belt. Therefore, the simulated seasonal variations are in better agreement with those observed in the rainfall belts in the mainland of China and the South China Sea.

From the analysis we can see that this model can simulate most features of the observation in the Asian summer monsoon region.

Results of Experiment on the Anomaly Snow Cover over the Tibetan Plateau

The anomaly snow cover over the Tibetan Plateau in winter will enhance the surface albedo in this area and make the surface heating effect decrease,

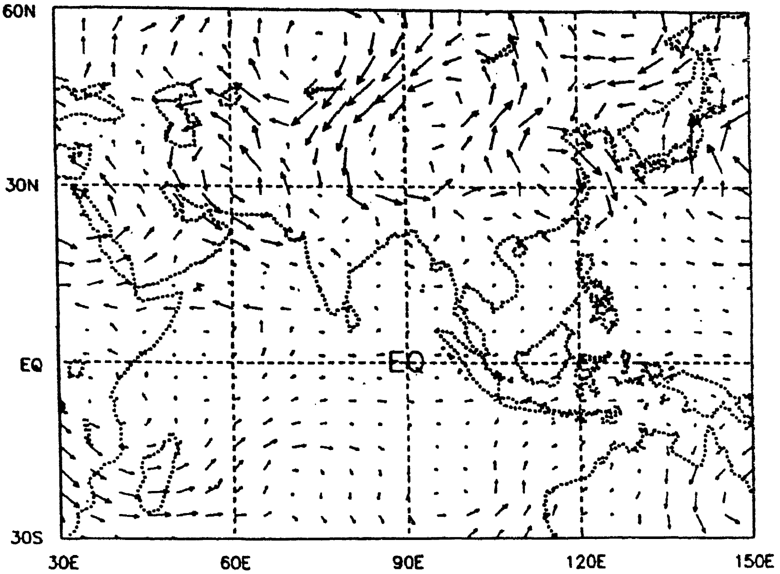


Fig 4(a) The difference of 250 hPa wind vectors in July

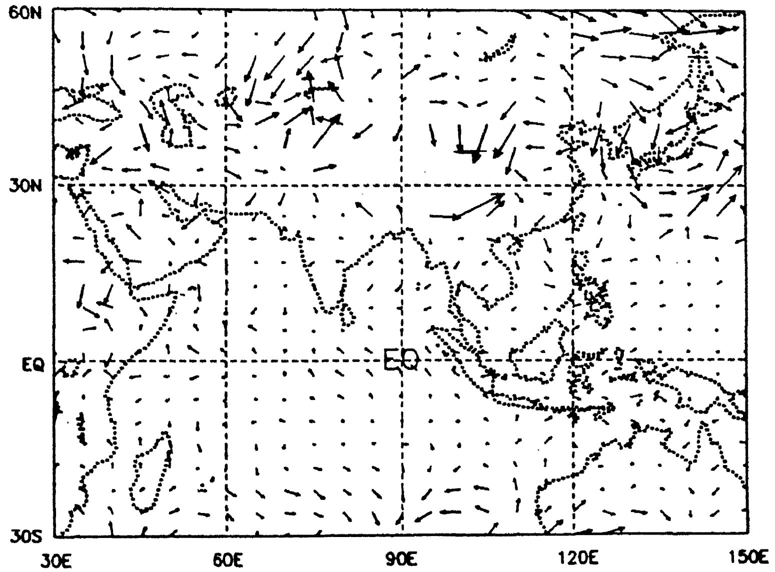


Fig 4(b) The difference of 975 hPa wind vectors in July

then cause the variation of circulation patterns. Now we want to test if this abnormal variation can influence the formation and evolution of the Asian summer monsoon circulation. Figs 4*a* and 4*b* are the differences of simulated monthly mean wind vectors at 250 hPa and 975 hPa level in July, respectively (Exp. 2 minus Exp. 1). We can see that at the upper level the anticyclone over the Plateau and west Asia has been weakened but it strengthened over East Japan. At the lower level, the SW monsoon flows in the Indian monsoon region and the South China Sea are weakened slightly, while a differential cyclonic circulation appears over the South-east Japan. From Figs 4*a* and 4*b* we can see that at the upper and lower level the differences of wind vectors are larger in middle and high latitudes in low latitudes and tropics.

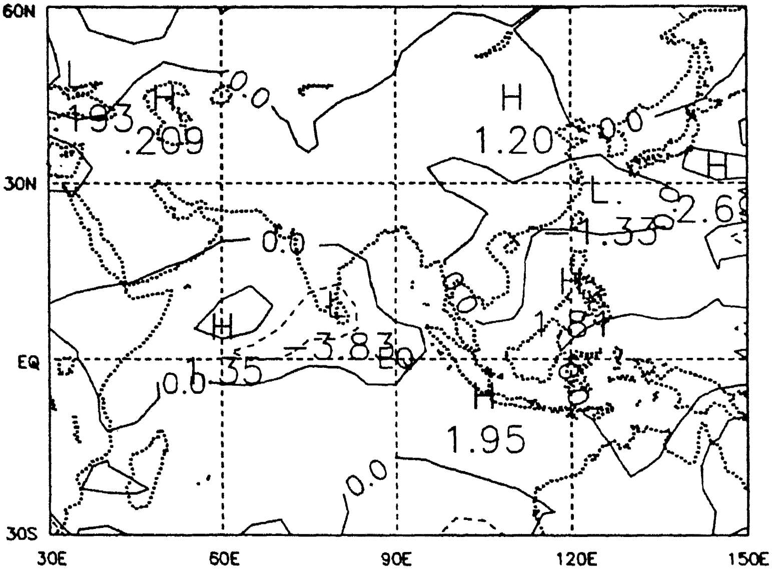


Fig 5 Difference of monthly mean precipitation in July (Exp. 2 minus Exp. 1).

The difference of the simulated monthly mean precipitation in July (Exp. 2 minus Exp. 1) is shown in Fig. 5. The observation show that if the snow cover over the Tibetan Plateau abnormally increased in winter, the precipitation in the middle and lower reaches of Yangze River will decrease and the mainland rainfall belt will shift southward in next summer[6]. From Fig 5, it is seen that the precipitation decreased in the middle and lower reaches of Yangze River and South India while it increased in North China, the South China Sea and East Japan. The simulated results are in closer agreement with those observed.

The result of the difference of the simulated monthly surface air temperature in July (Exp. 2 minus Exp. 1) shows that the surface air temperature is lower in most of China while it is higher in Indochina Peninsula and East Arabia Peninsula. The latitude-time cross sections of zonal wind and rainfall also show that the the polar jet over the north side of Plateau weakened and in east Asia the rainfall in Yangze River decreased but it increased in South China and South China Sea (Figs are omitted).

The above analysis shows that the response of the atmospheric general circulation to the abnormal snow cover over the Plateau is very sensitive, especially for the Indian summer monsoon system. Though this anomaly happens in winter, its effects can expend to the whole summer.

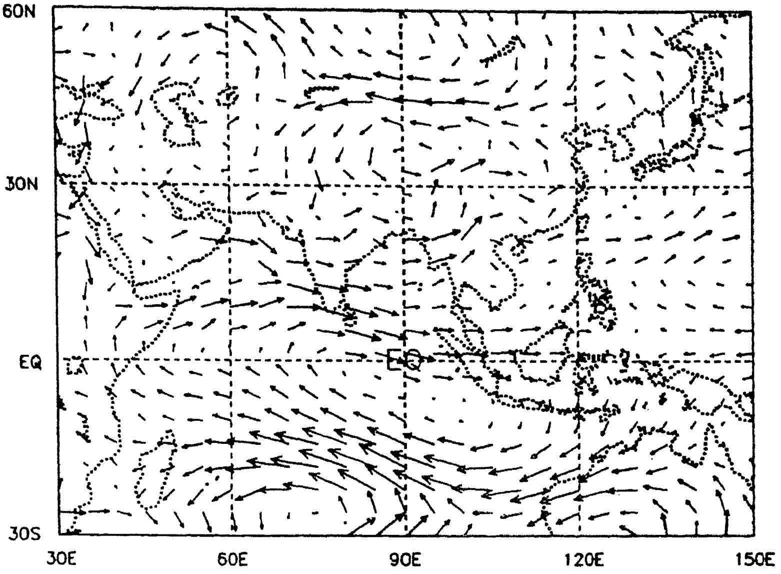


Fig 6(a) The difference of 250 hPa wind vectors in July

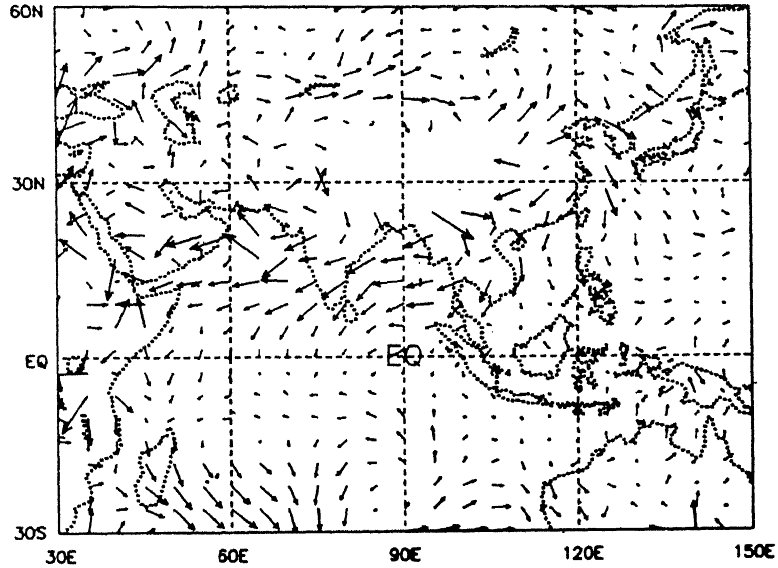


Fig 6(b) The difference of 975 hPa wind vectors in July

The atmospheric heat source anomaly over the Plateau in summer also has large effects on the circulation. The differences of the monthly mean wind vectors at 250 hPa and 975 hPa level in July are shown in Figs 6*a* and 6*b*, respectively (Exp. 3 minus Exp. 1). We can see that at the upper level, the South Asian anticyclone and the equatorial anticyclone are weakened obviously. Thus the upper easterly flows over the south Asia and cross-equatorial flows over the west of 90°E are also weakened significantly, while the anticyclone over the western Pacific is intensified clearly. At the lower level, there is a differential anticyclone from east Asia to west Asia. In the Indian summer monsoon area, the Somali jet and the SW monsoon flows over the Arabian Sea, the Indian subcontinent and Bay of Bengal are weakened obviously. The Indian monsoon trough disappeared and the cross-equatorial flows over the Indian Ocean are also weakened. In the east Asian monsoon area the SW monsoon flows over the East China and the north part of the South China Sea are weakened slightly. The south easterly flows over the western pacific are also weakened. From above analysis it is seen that the atmospheric diabatic heating has great influences on the Indian summer monsoon circulation, but for the east Asian summer monsoon system, its influence is not as obvious as that in the Indian monsoon system.

The difference of the simulated precipitation in July (Exp. 3 minus Exp. 1) is shown in Fig. 7. The figure shows that the precipitation increased in north India, south of Bay of Bengal and east of the South China Sea while it decreased in the Western Pacific, South India and East Arabia Peninsula. The results suggest that the atmospheric heating anomaly over the Plateau has important influence on the precipitation in low-latitude region of Asia.

The results of the latitude-time cross sections of the zonal wind at the upper level and the rainfall show that if there is no atmospheric diabatic heating over the Tibetan Plateau, the zonal wind speed over the Plateau and its neighbouring area will decrease obviously, its seasonal variation is not clear and at the upper level the easterly flows over the low-latitude of south Asia in summer disappear. In east Asia, the rainfall in the middle latitude of East China also decreased. In July the surface air temperature in most of Asia is lower, especially in East China and West India (Figs are omitted).

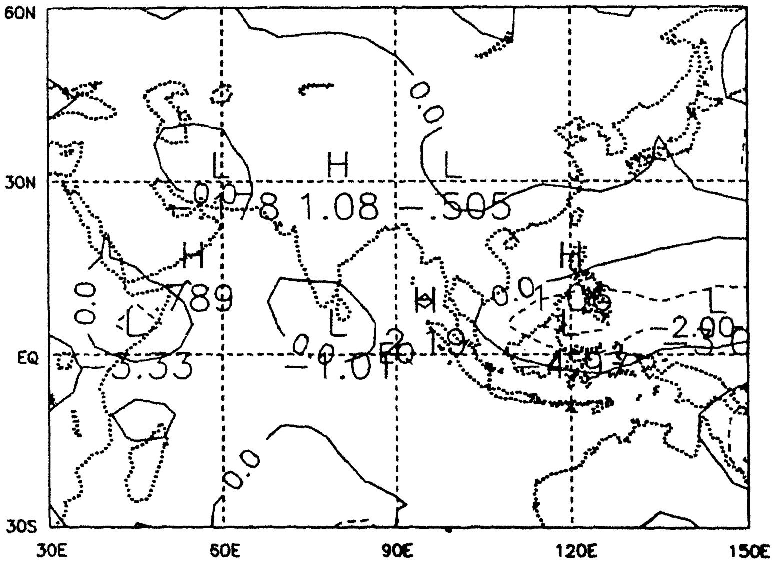


Fig 7 The difference of rainfall rate (mm/day) in July

The above discussion indicates that the atmospheric diabatic heating over the Tibetan Plateau is very important for the formation and evolution of Indian summer monsoon. Without this heating, the distribution of the heating fields over Asia will be changed and cause the variation of the Asian summer monsoon circulation: Indian summer monsoon disappear and the east Asian summer monsoon will be weakened.

Summary

Based on the above analysis and discussion, we can come to the following conclusions:

(1) The control experiment shows that the AGCM can reproduce most features of the observation in the Asian summer monsoon area.

(2) The effects of the anomaly snow cover over the Tibetan Plateau in winter can expend to the whole summer. A larger snow albedo than the normal on the Plateau in winter will weaken the summer monsoon in Asia. The precipitation in Yangze River will decrease and the rainfall belt in the mainland of China will shift southward which are in better agreement with those observed.

(3) The atmospheric diabatic heating over the Tibetan Plateau in summer is very important for the formation and evolution of the Asian summer monsoon circulation, especially for the Indian summer monsoon system. If this heating is removed, the Indian summer monsoon will disappear and the east Asian summer monsoon will be weakened.

References

1 A Kasahara and W M Washington *JAS* **28** (1971) 657-701
2 D G Hahn and S Manabe *JAS* **32** (1975) 1515-1541
3 L X Chen *et al.* *First Sino-American Mountain Meteorology* American Meteorological Society Boston (1982) 265-289
4 H L Kuo and Y F Qian *Mon Wea Rev* **110** (1982) 1879-1897
5 L X Chen *et al.* *Acta Met Sinica* 7 (1983) No 1, 31-49
6 L T Chen and Z X Yan *Collected Paers on the Medium-long Term Hydrometeorological Forecast (I)*, Water and Electricity Press Beijing China (in Chinese) (1979)

EAST ASIAN SUMMER MONSOON—A CASE STUDY

QING-CUN ZENG, BANG-LIN ZHANG, YOU-LIN LIANG AND
SI-XIONG ZHAO
Centre for Climate and Environment Prediction Research
Institute of Atmospheric Physics, Chinese Academy of Sciences
Beijing-100 080, *China*

East Asian summer monsoon for 1992 is analyzed. It is shown that the East Asian summer monsoon can be tracked back to the "southeast monsoon" (southeastly wind at surface and southwestly wind at 850hPa) created during spring and connected with the large seasonal variation of atmospheric general circulation over the tropical and subtropical western Pacific Ocean. During early and middle summer after the onset of Indian summer monsoon, the East Asian summer monsoon is modified and combined with the Indian summer monsoon as a whole southwestly wind system, but the influence of tropical and subtropical flow from western Pacific Ocean and South China Sea is also very clear. According to the similarity of patterns the division of seasons of atmospheric circulation also shows that the summer pattern comes into India and East Asia almost simultaneously. The Indian monsoon region and East Asian monsoon region are indeed characterized by large seasonality (normalized seasonal variation). The other regions with large seasonality are tropical western Pacific Ocean and Australian monsoon region (both are directly connected with East Asian monsoon region), as well as those separated regions which are coincided with the seasonal migration of subtropical high over Pacific Ocean in both Northern and Southern Hemispheres. The monsoon rainfall belt is along the so-called shear line which is an elongated narrow zone with large wind shear and possesses character of quasistationary front (Mei-yu front). Superimposed on the shear line there are mesoscale disturbances causing heavy precipitation.

Key Words: East Asian Summer Monsoon; Southwest Monsoon; Tropical & Sub-tropical Flow; Mesoscale Disturbances; Mei-yu Stage; Climate Dynamics

Introduction

Monsoon is a climatological concept which expresses the most significant seasonal variation of climate regime in the tropical-subtropical zone of planetary scale, although it is still difficult to give a precise definition. Once the monsoon is a tropical-subtropical phenomenon of planetary scale associated with the seasonal variation, it is not difficult to imagine that the monsoon is closely related or strongly influenced by the interaction of air flows in the two, Northern and Southern, Hemispheres, and by the continent-air-ocean interaction, because the first cause of seasonal cycle is due to the periodical annual variation of solar declination between Northern and Southern Regressions, and that there is large seasonal variation of thermal contrast between the continent and ocean. However, the problem on the nature, especially the cause of monsoon, is still a big argument in the present time due to its complexity and the lack of adequate observational data and methods for their analyses.

Indian monsoon, i.e. the monsoon over Indian Peninsula and the Indian Ocean, is the typical one, and is the first branch of the monsoon which has been investigated by means of modern meteorology. The origin of Indian summer monsoon, the propagation of cross equatorial current, the geometrical constraints of East Africa and Tibetan Plateau, the onset of summer monsoon and its relation to the starting of rainfall and to the seasonal transition of planetary scale atmospheric general circulation and so on have been well studied due to the good observation data, especially after MONEX.

It is well known that East Asian summer monsoon is that branch of monsoon which is far extended into the north even more than 45°N. However, due to its complexity of nature and the lack of adequate observational data over the vast region of western Pacific Ocean, this branch has not been well understood by the world community of meteorologists. Even in the present some meteorologists still think that the East Asian summer monsoon is only a direct extension of the Indian summer monsoon, i.e., the extension of south-west wind coming from the Indian Ocean and passing Indian Peninsula, although many Chinese, Japanese and other scientists, starting from Chu's classical work[1], have indicated, that the Mei-yu (summer monsoon rainfall in Yangtze River Valley) and Baiu (in Japan) are due to the confluence of south-west flow (tropical air mass), south-east flow (tropical-subtropical air mass) and north-west and north-east flow (cold air mass), and that the strongest winter monsoon, in contrary to the summer monsoon, can propagate from the East Asian continent into the region south to Philippines and even penetrate into Southern Hemisphere (Li, 1935)[2]. This means that at least one channel of air mass exchange between Northern and Southern Hemispheres is located over the equatorial western Pacific Ocean.

Indeed, in the last 30 years, after improvement of observation by the satellite in the vast region of western Pacific Ocean, especially the period of TOGA Program, and some intensive investigations, it is revealed that large part of air flow considered as summer monsoon and reaching the East Asian continent can be tracked back to the region covered South China Sea and the vicinity of Philippines and to the western Pacific Ocean, although some other part comes from Indian Ocean. There exists a negative correlation between the two branches of summer monsoon, the Indian monsoon and East Asian monsoon. The structure of East Asian monsoon is more complicated and variable due to the strong interaction of the three flows mentioned above, and that there may be more transport of heat and other elements in that region and, consequently, may have significantly mutual influence with the variability of global general circulation. For example, some preliminary studies have shown that there is interaction of winter-summer monsoons and the ENSO. It is expected that the study on the East Asian monsoon, its relationship to the Indian monsoon, other branches of monsoon, and the global general circulation will be intensified.

In this paper, the authors give a case study (year 1992) in order to demonstrate clearly some basic pictures and characteristics of East Asian monsoon, sch as the seasonality (measure of seasonal variation), division of seasons, typical patterns of summer monsoon circulation and the associated rainfall belts during different stages of their evolution, the vertical structure of summer monsoon, and the role of weather and mesoscale systems in the maintenance of large and planetary scale monsoon flow.

The Seasonality

Monsoon climate is characterized by a large amplitude (variability) of seasonal cycle of wind, rainfall, and, may be, other meteorological elements. We first define the seasonality of a function or vector $F(\theta, \lambda, p, t)$ varying with time t and defined at the geographic point (θ,λ) and on a given p isopressure surface as follows:-

Let F_w and F_s are the typical fields of F for winter time and summer time respectively. For example, F_w and F_s are taken as January and July monthly means respectively. Obviously, $\Delta = |F_w - F_s|$ is the "amplitude" of seasonal cycle if F_w *and* F_s are indeed two extreme regimes of F. Let $\overline{F}$ be the annual mean of F, or simplify taken as $(F_w + F_s)/2$, the seasonality δ of F for a given area S can be defined as the normalized variability.

$$\delta \equiv \frac{\|\Delta\|}{\|\overline{F}\|}, \qquad \ldots (1)$$

where $\|\bullet\|$ is the norm, for example,

$$\|F\|^2 = \frac{1}{S}\int\int_S |F|^2 a^2 \sin\theta d\theta d\lambda, \qquad \ldots (2)$$

a is the radius of the Earth, and S denote the area and its measure both. S can also denote a point (θ, λ), if $\overline{F}(\theta, \lambda, p)$ is different from zero.

Figs 1-3 represent the January mean wind $\mathbf{v}_w$, July mean $\mathbf{v}_s$, and $\mathbf{v}_s - \mathbf{v}_w$ at 850hPa for 1992 respectively. Fig. 4 is the seasonality δ of wind at 850hPa for 1992 and is calculated for every grid point in the latitudinal zone 40°S-40°N. Note that in Fig. 4 we take $\bar{\mathbf{v}} \equiv (\mathbf{v}_s + \mathbf{v}_w)/2$, and that the annual mean wind field is similar to it, hence the associated δ pattern (not shown here) is also similar to Fig. 4. It can be seen that in all maps of $\mathbf{v}_w$, $\mathbf{v}_s$ *and* $\mathbf{v}_s - \mathbf{v}_w$ the meridional

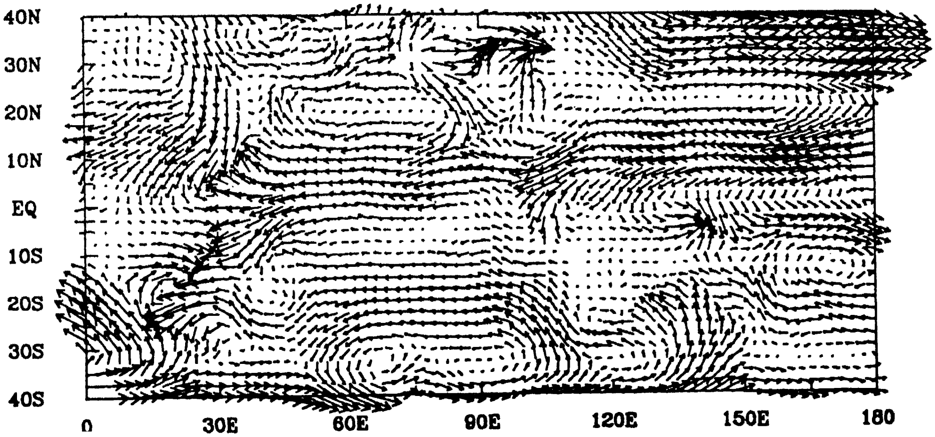


Fig 1 January mean wind $\mathbf{v}_w$ at 850hPa for 1992.

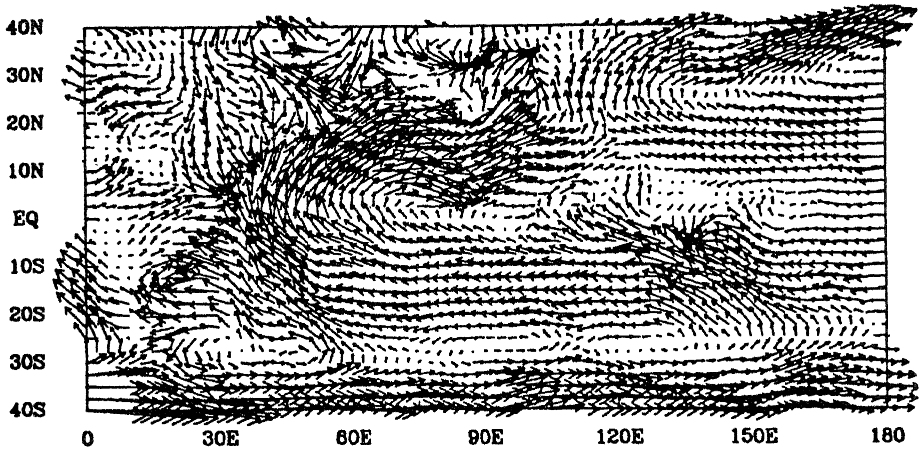


Fig 2 July mean wind $\mathbf{v}_s$ at 850hPa for 1992

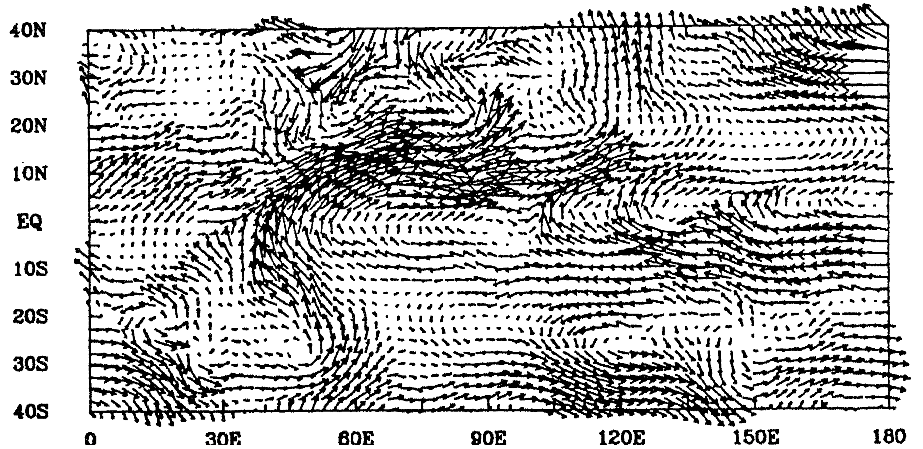


Fig 3 The difference of July from January means winds $\mathbf{v}_s - \mathbf{v}_w$ for 1992.

component is significant in the Indian monsoon region and East Asian monsoon region where the Δ and δ are also pronounced, and that the East Asian monsoon region is connected (extended) to a region with equally large or even larger Δ and δ which is located from 125°E to 165°W and from 15°S to 10°N (not uniform) and is correspondent to the Australian monsoon region and ITCZ with pronounced seasonal cycle in the Pacific Ocean. In other place of tropical zone both Δ and δ are small or not competitive even in the African monsoon and South American monsoon regions. However, in association with the seasonal migration of subtropical high over Pacific Ocean in both the Northern and Southern Hemispheres where Δ and δ are also pronounced, although they are not considered as monsoon regions in the literature.

It is clear in Figs 1-3, there are two channels of cross-equatorial flow, one (Channel 1) located in the Indian Ocean and associated with Indian monsoon, the other (Channel 2) located in the equatorial region near South China Sea and Philippines and connected with East Asian monsoon. Channel 1 is very intensive during summer, while Channel 2 is clear both in summer and winter, especially during winter. Due to the large difference $\mathbf{v}_s-\mathbf{v}_w$ in Channel 2 there δ is also large (not uniform).

One might also consider the three dimensional structure of seasonality of atmospheric general circulation by computing $\Delta(\theta, \lambda, p)$ and $\delta(\theta, \lambda, p)$.

Seasonal Transition of Atmospheric General Circulation

The seasonal transition and division of seasons of atmospheric general circulation can be determined by the following method (Zeng and Zhang, 1991)[3].

Denoting $\bar{F}\equiv(F_s+F_w)/2$, we calculated the depature $F'(\theta, \lambda, p, t)\equiv F(\theta, \lambda, p, t)-\bar{F}(\theta, \lambda, p)$, $F'_w\equiv F_w-F$, and $F'_s\equiv F_s-\bar{F}$. Obviously, $F'_s=F'_w$. The simi-

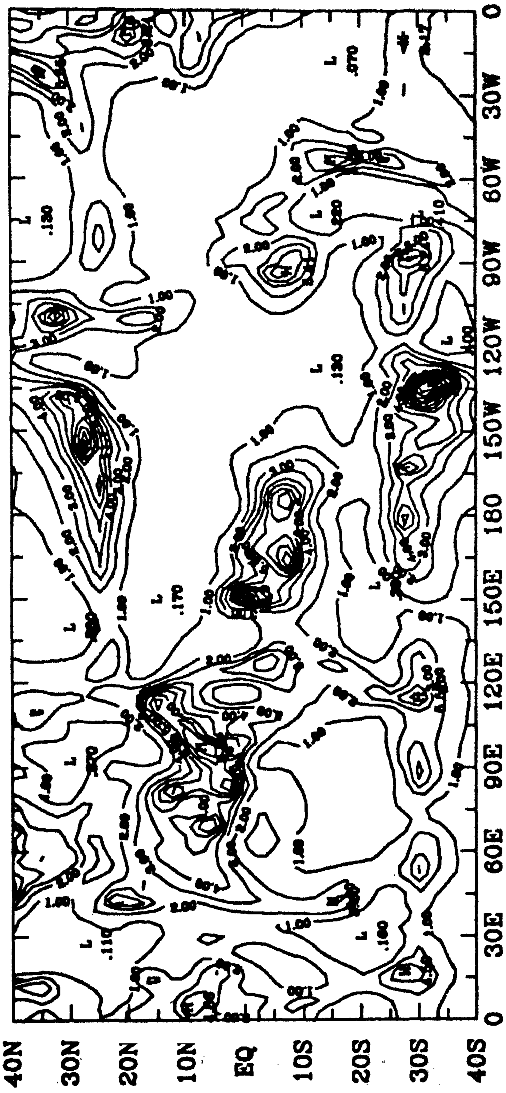


Fig 4 The seasonality δ (see (1)) of wind at 850hPa for 1992.

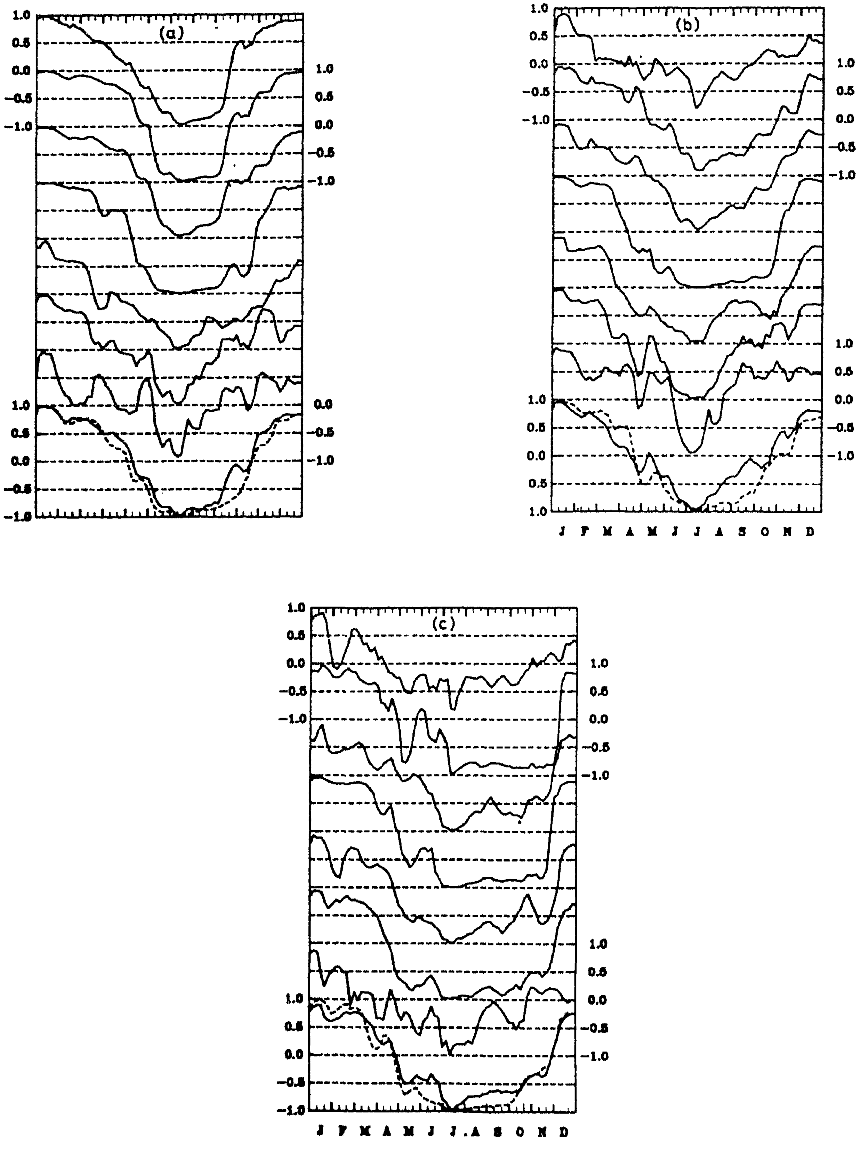


Fig 5 The curves of $R(t)$ which defines the division of seasons for different areas. (*a*) 60°E-100°E, (*b*) 100°E-140°E, (*c*) 140°E-160°E. From above to below: 40°N-20°N, 30°N-10°N, 20°N-0°N, 10°N-10°S, 0°-20°S, 10°S-30°S, 20°S-40°S, and 40°N-40°S respectively. The solid line for wind at 850 hPa, and the dashed for wind at 300hPa.

larity of F'_w and $F'(\theta, \lambda, p, t)$ for a given area S is defined as the normalized inner production

$$R(t; p) = (F', F'_w)/[\|F'\| \cdot \|F'_w\|], \qquad \ldots (3)$$

where the inner production (F', F'_w) is given by

$$(F', F'_w) \equiv \frac{1}{S}\iint_S F' \cdot F'_w a^2 \sin\theta \mathrm{d}\theta \mathrm{d}\lambda \qquad \ldots (4a)$$

if F is a scalar function, or by

$$(F', F'_w) \equiv \frac{1}{S}\iint_S \mathbf{F}' \cdot \mathbf{F}'_w a^2 \sin\theta \mathrm{d}\theta \mathrm{d}\lambda, \qquad \ldots (4b)$$

if F represents a vector, for example, the wind **v**.

It is clear, that typical winter regime is identical to $R=1$, and typical summer regime is $R=-1$. We can divide seasons as follows: $0.5<R\leqslant 1$ – winter, $-1\leqslant R<-0.5$ – summer, $-0.5<R<0.5$ – transitional seasons (spring or autumn). Note, precisely speaking, the regime $-0.5<R<0.5$ is correspondent to autumn only if F_w is taken as the typical field for the following winter. Besides, the ratio T_{sp}/T_w is a measure of the abruptness of transition from winter to summer, where T_{sp} and T_w are the duration times of spring and winter respectively. ively.

Taking F as the 15 day running mean of wind **v** for 1992, Fig. 5 shows the $R(t, p)$ at different levels p(mostly $p=850$hPa, and some for $p=300$hPa) and for different areas S from 60°E to 160°E and from 40°S to 40°N with interval 40° along the latitudinal cycle (except for 140°E-160°E) and 20° along the meridian. It can be seen, that according to our definition of seasons of atmospheric general circulation, in the wind field at 850hPa we have (1) Spring comes ($R(t)=0.5$) earlier on about March 15 (from February 10 to April 10 in different subsector) in the sector (100°E-140°E, 40°S-40°N) where the East Asian monsoon region is located, on about March 25 in the sector (140°E-160°E, 40°S-40°N) but with earliest date, February 1, in the subsector (100°E-140°E, 20°S-0°S), and on about April 10 in the sector (60°E-100°E) where the Indian monsoon region is located with latest day, May 5 to 10 in the subsectors (60°E-100°E, 0°N-20°N) and (60°E-100°E, 10°N-30°N); (2) Summer comes ($R(t)=-0.5$) earliest from April 20 to May 10 in the subsector (140°-160°E, 0°S – 30°S), propagates north-westly into the subsectors (100°-140°E, 20°S-10°N) and (60°E-100°E, 10°S-10°N) on about May 15, and then northwards into 30°N, China and India, simultaneously on about June 10 and into North China, Japan and northeast area about July 10; (3) in the Northern Hemisphere autumn comes ($R(t)=-0.5$ after the summer) in the sector (60°E-100°E) earlier on about September 10, in the sector (100°E-140°E) on about September 25 (from September 20 in the north to October 25 along Equator) with exception for the subsector (100°E-140°E, 20°N-40°N) where the summer lasts for a very short time in 1992 and ended on July 25, and in the sector (140°E-160°E) much later than October 25 (from October 25 to December 5) except for the subsector (140°E-160°E, 20°N-40°N) where the summer is ended also early. Besides, the seasonal transitions are not synchronic in the vertical, for example, summer comes much earlier at 300hPa than at 850hPa, but is ended on the same time or even later. Note, at 300hPa summer comes as early as on May 1 in the sector (140°-160°E, 40°S-40°N).

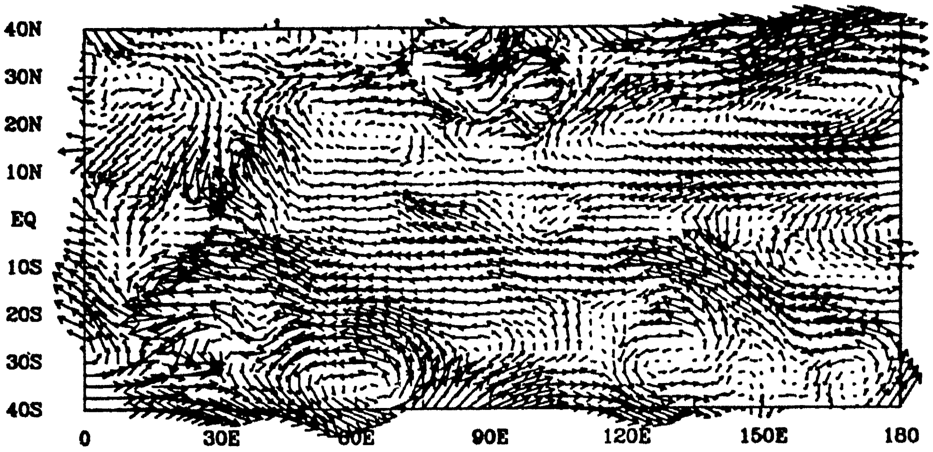


Fig 6*a* The mean wind field at 850hPa for April 16-30, 1992.

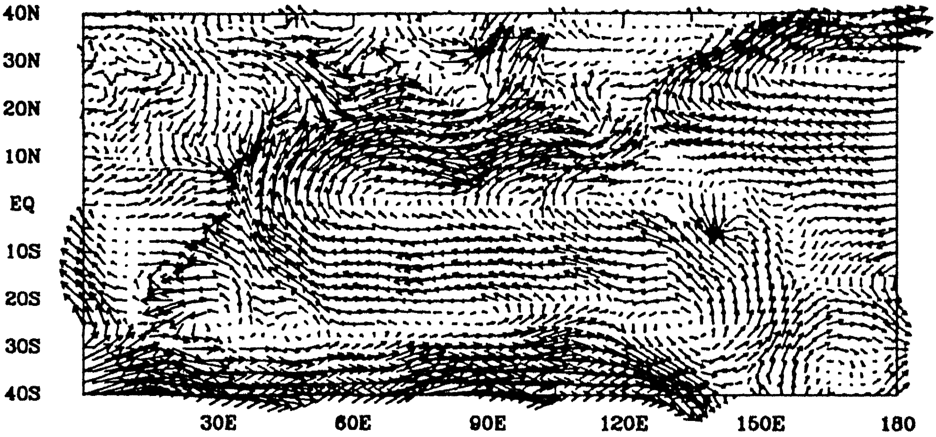


Fig 6*b* The same as Fig. 6*a* but for June 16-30.

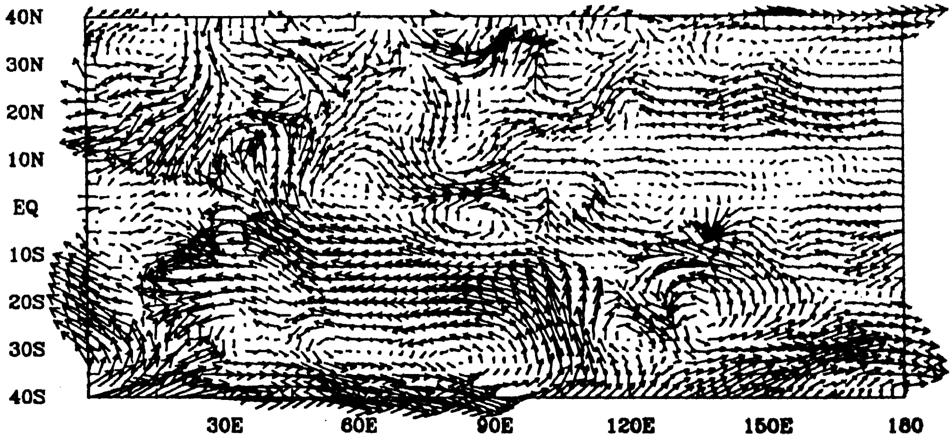


Fig 6*c* The same as in Fig. 6*a* but for September 16-30.

According to the results described in these two sections, it can be imagined that East Asian summer monsoon is more closely related to the seasonal variation of atmospheric generation circulation on the western Pacific Ocean, although connected and strengthened by the Indian summer monsoon.

Note, that the onset and the northern boundary of summer monsoon were determined by other criteria, for example by surface θ_{sw} (Tu and Hwang, 1944)[4] or by the so-called monsoon index (Tao *et al.*, 1987, 1988)[5,6] and that the division of seasons of general circulation and the abrupt seasonal transition were determined by the jet location and other characteristics[7-9]. Our new results are agreed with the previous ones to a large extent. May be due to the EI Nino or other interannual variations in 1992 spring comes too early and lasts for too long time, the transition from winter to spring over East Asia is not abrupt, although from $R=0$ to $R=-0.5$ is more or less abruptly. However the transition from summer to winter in 140°E-160°E region is very abrupt.

Different Stages of East Asian Monsoon

According to the division of seasons by our new method we can also divide the East Asian monsoon into several stages which are agreed to some extent with the definition of common sense. Obviously, the typical summer monsoon regime and typical winter monsoon regime are identical to $R=-1$ and $R=1$ and represented by the July mean and January mean respectively. The typical spring and autumn are identical to the first event with $R=0$ and last event with $R=0$ and given by April 16-30 and September 16-30 respectively for 1992 as can be seen in Fig. 6. As here we are mostly interested in the summer monsoon we will discuss summer season and the adjacent seasons in some detail.

(*a*) *Southeast Monsoon Stage*

During the whole spring and early summer, lasted from late March to the beginning of Mei-Yu season in the Middle June, there are southeasterly wind (so-called southeast monsoon in Chinese literature) at the surface and southwesterly wind at 850hPa in the southeast part of China and the adjacent area over ocean. This is the west wing of big anticyclonic gyre of wind system over the vast region of western Pacific Ocean. The strong winter monsoon over South China Sea completely ceases. The precipitation region occupies south China and extended to the western Pacific Ocean. The other rainfall belt is associated with ITCZ and its extension, located in Thailand, Malaysia, Philippines, Indonesia and their vicinity. The character of the precipitation is of quasi-stationary front type in the former region but of convective type in the later. Fig **6a** and Fig. **7a** are the typical patterns of 850hPa wind and precipitation. These figures are obtained by averaging for April 16-30, 1992 and associated with $R=0$. It can be seen that the southwest wind at 850hPa over southeast China is closely connected with easterly wind over tropical western Pacific Ocean, but clearly separated from the cross equatorial flow over Indian Ocean. The late just begins in the later April and is limited in a localized area along the Equator.

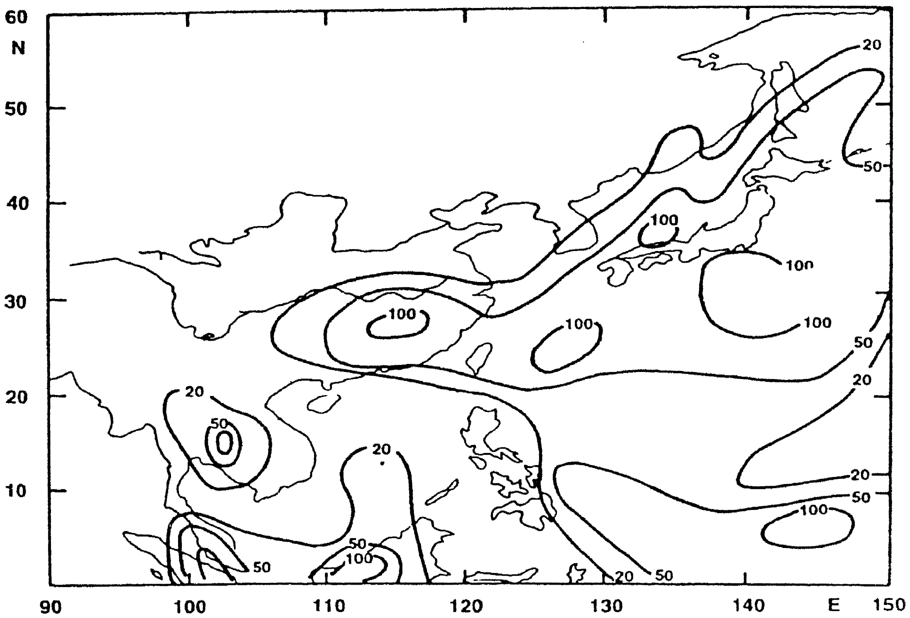


Fig 7*a* The accumulated precipitation for April 16-30, 1992. The unit is mm.

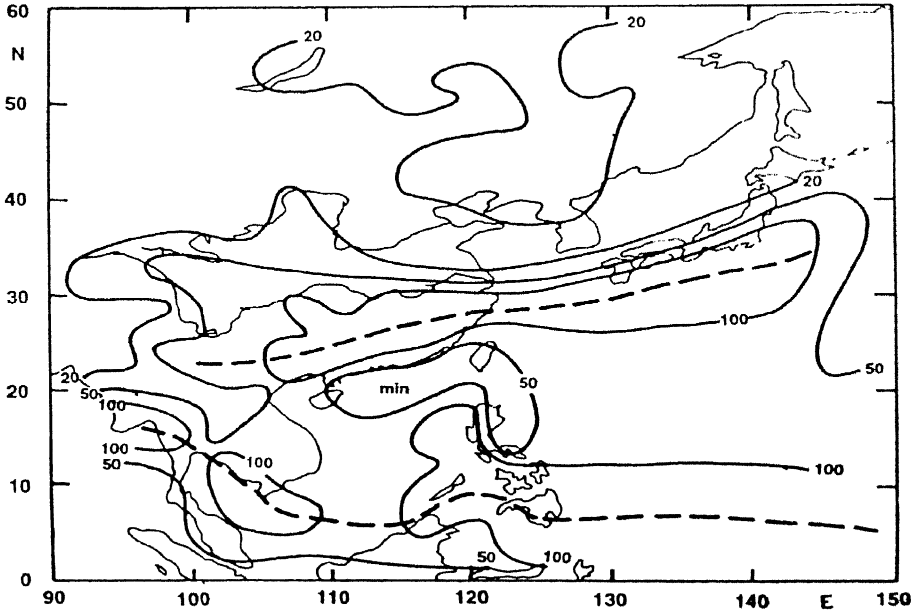


Fig 7*b* The same as Fig. 7*a* but for June 16-30. The belt with maximum precipitation is indicated by heavy dashed line.

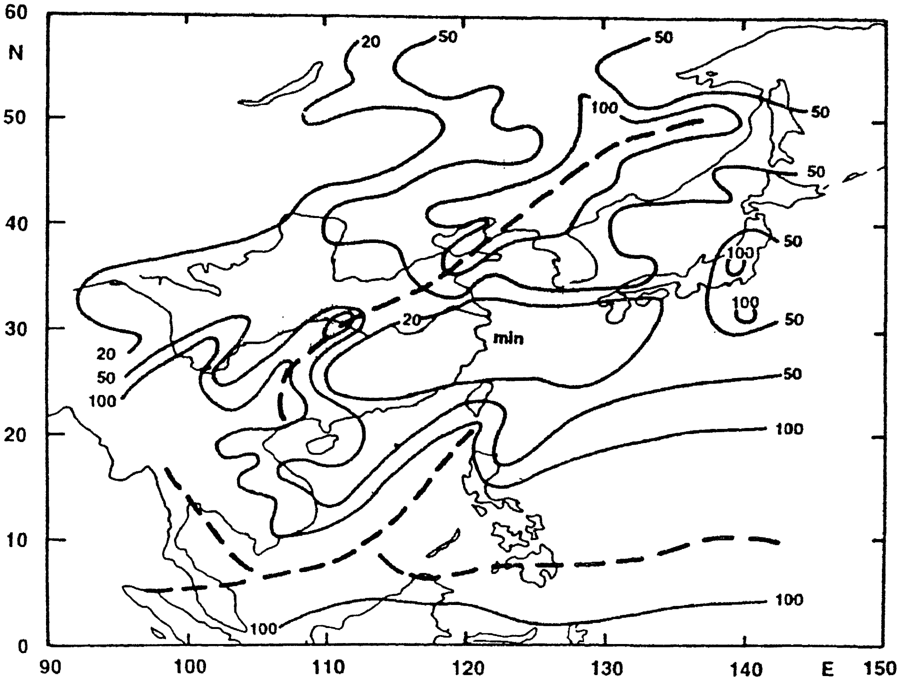


Fig 7*c* the same as Fig 7*b* but for July 16-30.

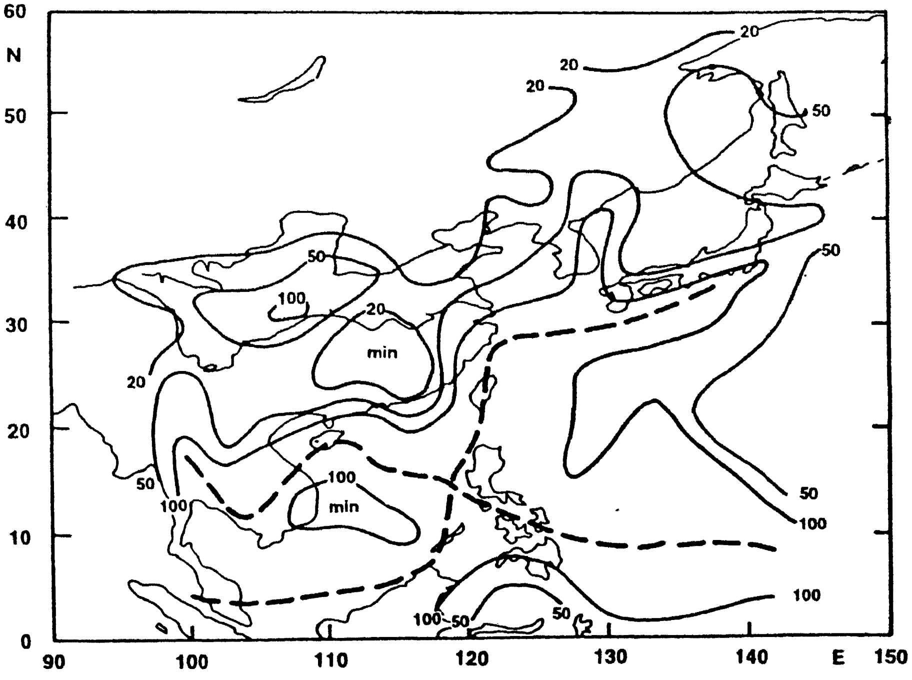


Fig 7*d* The same as Fig. 7*b* but for September 16-30.

(*b*) *Mei-yu Stage*

The southwestly wind at 850hPa over south China (hereafter referred as wind system A) is gradually intensified, extended northeastward, southward and southwestward and occupies very large region. During June 1-15, Indian monsoon (wind system I) comes into South India and is extended into Bay of Bengal. At the same time the cross equatorial southly wind frequently appears over South China Sea and the vicinity, and the ITCZ moves to the Northern Hemisphere. All these strengthen wind system A and make it connected with wind system I, and result in a intensive southwestly wind belt extended to Somalian jet Associated with wind system A there is southeasterly or southwesterly wind at the surface. Starting from Middle June and lasted for about 25 day or so, the quasistationary front migrates along Yangtze River to Hai River (30°N approximately) and extends to Japan in normal year but slightly south to them in 1992 (may be due to the influence of El Nino), and results in a bulk precipitation (Mei-Yu) there, Fig. 6*b* and Fig. 7*b* are the typical patterns obtained for June 16-30. Figs 8-10 show one example of synoptic analysis during Mei-yu period.

(*c*) *Hot Summer Stage*

After Mei-yu period the subtropical high over western pacific Ocean moves northward, its ridge is extended into East China along 30°N, and the ITCZ moves into 10°-20°N in the region (120°E-150°E) of western pacific Ocean. In association there is strong southwesterly wind system (system B) in the region centered at (130°E, 45°N)—Northeast China, south to it there is southeasterly wind, and over whole south China (south to 30°N in the normal year, but slightly more south in 1992) southwestly wind is prevailing from surface to the level higher than 850hPa (the wind field is not given here, it is similar to Fig. 2). The precipitation region jumps to North China or even Northeast China, Mei-yu is stopped, and hot weather occupies south China (see Fig. 7c), except for the case that typhoon is formed in the ITCZ and moves into mainland.

(*d*) *Autumn with Clear Sky Stage*

In August the subtropic high as well as the major precipitation region moves to or remains at their extreme north location, and due to ITCZ activities including typhoons there is another but separated precipitation zone in the tropic-subtropic zone over western Pacific Ocean and extended to the coast region of southeast China. Starting from September the summer monsoon regresses southward and weakens down, cool air mass with westerly-to-northwesterly wind component at the surface and in the lower troposphere frequently moves from the north to East Asian continent and completely occupies it in the late September. There is cool weather with clear sky and no or very few precipitation. Fig. 6*c* and Fig. 7*d* are the typical patterns (September 16-30).

The Structure of Mei-yu Weather System

The synoptic structures of East Asian summer monsoon have been thoroughly analyzed by Hsieh *et al.*[10,11] and many Chinese meteorologists. It is revealed

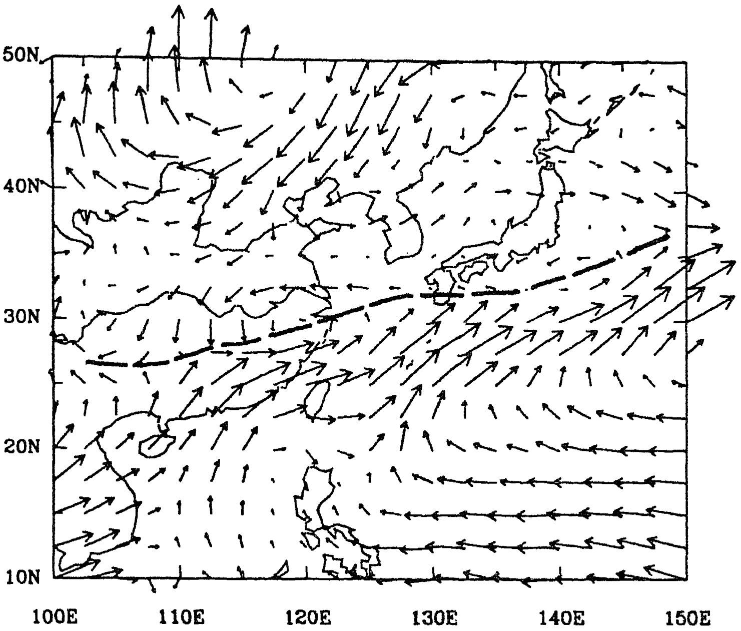


Fig 8 The wind field at 850hPa for 12Z, 16 June, 1992. The heavy dashed is the shear line.

that the structure of Mei-yu weather system is very unique. A typical example is given by Figs. 8-10 (for June 16, 1992), although the quasistationary front is located slightly south to the mean case.

It can be seen at 850hPa wind field (Fig. 8). There is a very narrow zone with very large wind shear and large convergence. This is called as shear line in Chinese literature. It is elongated from Japan vicinity to Southwest China and with almost east-westwards orientation. South to this line there is southwesterly wind system, whose western part comes from Indian Ocean and South China Sea area and is of tropical origin, but the eastern part can be tracked back to the easterly wind in the south and southwest sides in the subtropic region of western Pacific Ocean. North to the shear line there are southeasterly wind in some part and wind with northly component in other part, and at surface this wind system clearly comes from the cool air mass (it can be seen also in Fig. 9).

Fig. 9 shows the vertical structure across Mei-yu front. It is seen that the θ_{se} isolines are very dense and very cruted in the vicinity of Mei-yu front (shear line) due to the large contract of moisture, although the horizontal temperature gradient is only moderate or weak. In the lower troposphere, south to the front this is tongue of maximum θ_{se} representing the very wet summer monsoon air mass, but north to the front this is a shallow shield of minimum θ_{se} with northly wind representing the dry air mass of middle-high latitude origin. In both sides the upper troposphere is dry due to the lack of ascending motion except for the direct vicinity of Mei-yu front and the convection area in the south.

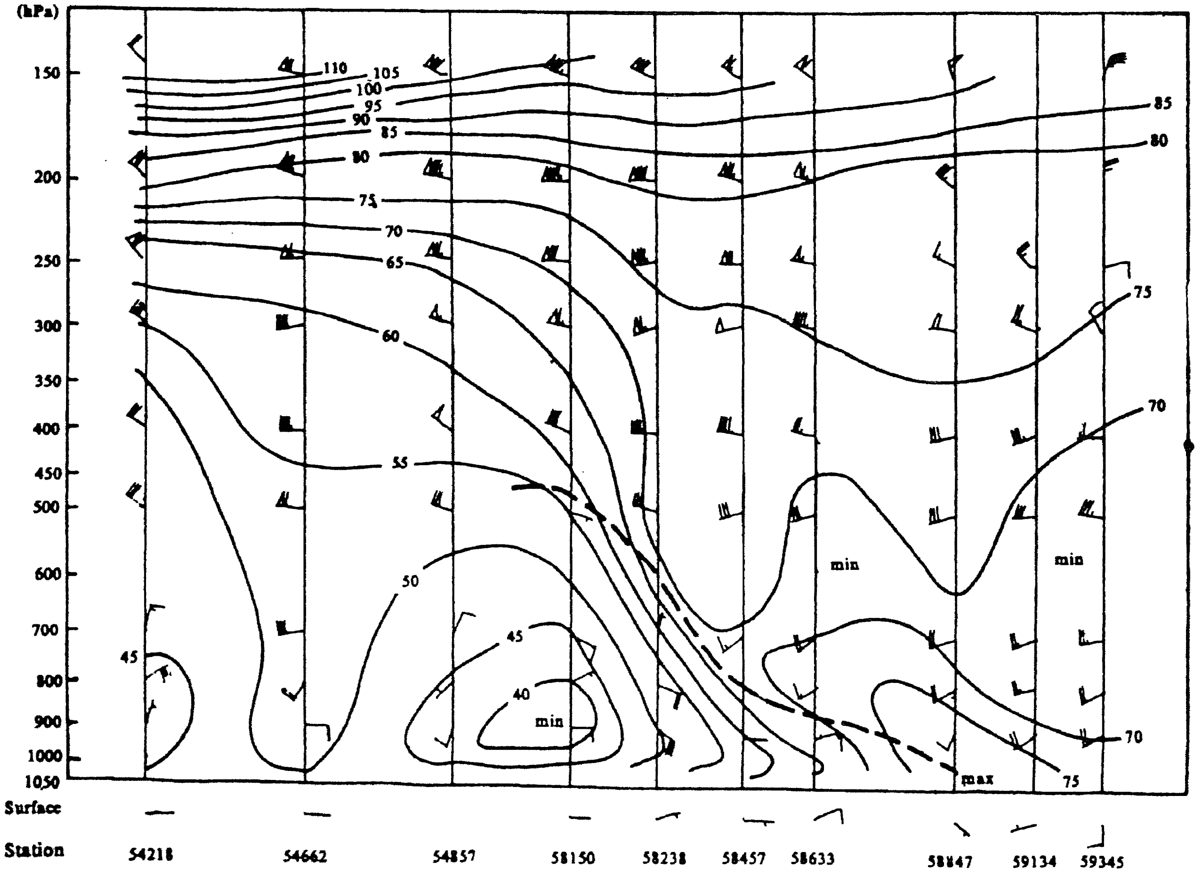


Fig 9 The vertical cross section along 120°E for 00Z, 16 June, 1992. The solid lines for θ_{se} and the heavy dashed indicates the shear line.

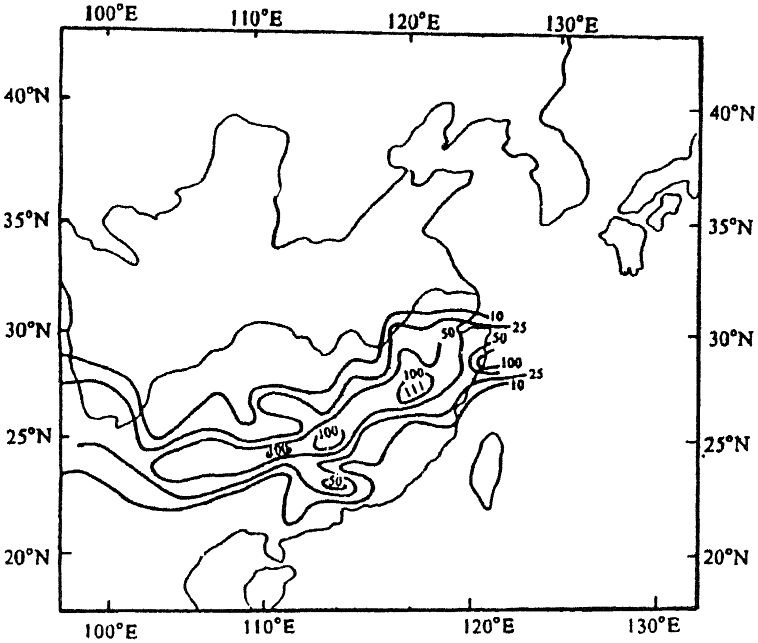


Fig 10 24 hour precipitation for 16-17, June, 1992. The unit is mm.

Fig. 10 shows the 24 hour accumulated precipitation. The region with heavy precipitation is just along the shear line, but there are a lot of mesoscale extreme centers representing the mesoscale feature. It is also revealed that the mesoscale and synoptic scale systems are very important in the maintenance of large scale monsoon flow as well as its super-geostrophic feature due to the eddy transport of momentum and energy (see also Zhao, 1988)[12]. Note the release of latent heat is also very important during the onset of summer monsoon as it has been emphasized by Krishnamurti and Ramanathan[13] and Zeng *et al.*[9].

Summary

Based on the analysis of observational data for 1992, the East Asian summer monsoon can be tracked back to the cease of Asian winter monsoon (northeastly wind at the surface and in the lower troposphere), to the creation of "Southeast monsoon" (southeasterly wind at the surface but southwesterly wind in the lower troposphere) during spring in the coastal region of southeast China and the adjacent ocean due to the large seasonal variation of atmospheric general circulation over the South China Sea and the tropical and subtropical western Pacific Ocean. The "southeast monsoon" is gradually extended into larger area, modified and intensified during early and middle summer by the Indian summer monsoon, cross equatorial flow over South China Sea and the vicinity of Philippines, and the subtropic flow due to the intensification and northwestward marching of subtropic high over Pacific Ocean in the Northern Hemisphere. After middle summer the East Asian summer monsoon reaches even 50°N, but its northern part is clearly connected with the westerly wind of subtropic high.

Based on the similarity of patterns, the division of seasons of atmospheric general circulation shows that the spring and summer come first in the tropical region of western Pacific Ocean, propagate northwestward into South China Sea and the vicinity. The summer (i.e., typical summer monsoon) comes to India and East Asia almost simultaneously.

There are two channels of cross equatorial air flow in the lower troposphere. One is located over the Indian Ocean and associated with Indian summer monsoon, the other is located over South China Sea and the vicinity of Philippines and associated with Asian winter monsoon and also influencing East Asian summer monsoon.

According to the computed seasonality (normalized seasonal variation), the Indian monsoon region and the East Asian monsoon region do indeed possess large seasonality. The East Asian monsoon region is closely connected with other regions of large seasonality, which are located in the tropical western Pacific Ocean and Australian monsoon region.

East Asian summer monsoon brings bulk precipitation in a large area, although nonuniformly with mesoscale structure. During the raining season (Mei-yu) along about 30°N in the East Asia the precipitation in along the shear line which is an elongated narrow zone with large wind shear due to the confluence of southeasterly wind from tropics, southeasterly wind from tropical and subtropical western Pacific Ocean, and northly wind from middle-high latitude in the lower troposphere. The shear line is associated with the Mei-yu quasistationary front with large contract of moisture in the cross section.

Acknowledgement

This work was supported by the Grant of China National Programme on Basic Research "Climate Dynamics and Prediction Theory".

References

1 C C Chu *J Chinese geogr Soc* **1** (1934) 1-27 [The English Translation was published in *Collected Scientific Papers, Meteorology* 1919-1949, Academia Sinica (1954) 475-493]

2 S Z Li Die kälteeinbruche in Ostasien Von *Inaugural Dissertation*, Univ Berlin/[The reprint was published in Collected Scientific Papers, Meteorology, 1919-1949, Academic Sinica (1954) 15-99]

3 Q C Zeng and B L Zhang *Chinese J atmos Sci* **16** (1992) 641-648

4 C W Tu and S S Hwang *Met Mag* **18** (1944) 81-92 [the English translation was published in *Collected Scientific Papers, Meteorology* (1954) 1919-1949, Academia Sinica, 519-534]

5 S Y Tao and L X Chen *In: Monsoon Meteorology* (*Eds* C P Chang and T N Krishnamurti) Oxford Univ Press (1987) 60-92

6 S Y Tao, W M Zhu and W Zhao *Scient Atmos Sinica Sp. Issue* [The English translation was published in Frontiers in Atmospheric Sciences (1992) 18-28 Allerton Press Inc]

7 T C Yeh, S Y Tao and M C Li *In: The Atmosphere and the Sea in Motion* (1959) 249-276

8 Y S Kao *Acta Meteorologica Sinica* **29** (1958) 83-92

9 Q C Zeng, X Z Liang and M H Zhang *Scient Atmos Sinica Sp Issue* [The English translation was published in Frontiers in Atmospheric Sciences (1992) 29-57 Alterton Press Inc.]

10 Y P Hsieh *Acta Meteorologica Sinica* **27** (1956) 1-24

11 Y P Hsieh and C T Tseng (Q.C. Zeng) *Tech Rep No.* 3 Central Meteorological Bureau of China (1957).

12 S X Zhao *Scient Atmos Sinica Sp Issue* [The English translation was published in Frontiers in Atmospheric Sciences (1992) 239-248 Alterton Press Inc.]

13 T N Krishnamurti and Y Ramanathan *J atmos Sci* **39** (1982) 1290-1306

A BRIEF REVIEW OF THE STUDIES ON THE SUMMER MONSOON IN CHINA

GUO XIAORONG
State Meteorological Administration, Beijing, China

(*Received and Accepted* 1 *October* 1993)

China is greatly affected by monsoon. Every year, the summer monsoon brings about abundant precipitation in China. Like India, the activity and variability of the summer monsoon is intimately related to the crop yields. Therefore, the Chinese meteorologists have for a long time been engaged in the numerous studies on the impact of the monsoon on the weather and climate in China, especially the droughts and floods, and obtained many important results. Owing to the limitation of space, it is impossible to make a comprehensive review of these achievements in this short report. This paper only summarizes several aspects studied on the summer monsoon in China during the past years.

Key Words: China-Summer Monsoon; Crop Yields; Main Rain-Bands; Somali Jet; South West Low Level Jet; Weather Systems

The Origins of the Summer Monsoon in China

The summer monsoon affecting China originates from three airflows: (*a*) the Indian summer monsoon—When it moves northward, mainly during the break monsoon period, the southwest monsoon may often advance deep into the China main; (*b*) the cross-equatorial airflow flowing over southeast Asia and the South China Sea—This is a kind of low level southwest airflow. Recent studies have shown that this airflow may be an independent summer monsoon system, although it may interact with the Indian summer monsoon. This airflow originates in Australia of the southern hemisphere and then streams across the equatorial region, mainly over Singapore and the North Borneo region. This cross-equatorial air-current is considered to be another significant exchange channel of air mass between the two hemispheres. It can flow northward over South China, even further over the reaches of the Yangtze River; (*c*) the southeast monsoon coming from the southern flank of the subtropical high over the West Pacific—This airflow may turn into the southwest airflow when it flows along the western side of this height. Sometimes one branch of the above three monsoon airflows exists alone, sometimes their combination can be observed over a certain area in China. Therefore, the activity of the summer monsoon in China is much more complicated than in other monsoon regions.

The Advance and Retreat of the Summer Monsoon in China and Its Relationship to the Movement of the Main Rain-Bands in China

The advance and retreat of the summer monsoon in China is in stepwise way instead of continuous way. The distribution of the heavy rainfalls in China during 1953-77 shows four belts of heavy rainfall with E-W orientation, located in South China, the valley of Yantze River, the reaches of the Yellow River and

North China respectively. This characteristic distribution of heavy rainfalls mainly corresponds to the four relative stable stages and three abrupt northward advances of the summer monsoon in China. Therefore, the activity of the summer monsoon will strongly influence the distribution and shift of the precipitation belts in China. It is found that the drought and flood happened in the south of China are associated with the anomalous variation of the South Asian monsoon circulation.

The Somali Jet and Its Relationship to the Southwest Low Level Jet

The Somali jet is the most stable and deepest cross-equatorial airflow in the Asian monsoon region. There is a very good correlation between it and the southwest monsoon located in Southeast China. Its variation and orientation make an impact on the southwest low level jet, that originates from the Bay of Bengal and South China Sea. Some numerical simulations have been made with the Somali jet The results can be summarized as follows:-

(*a* The non-adiabatic heating plays an important role in maintaining and evolving the Somali jet and its adjacent circulation. The non-adiabatic heating is controlled mainly by two factors. One is the non-homogeneity of the underlying surface such as the difference between land and ocean, that decides the character of the planetary boundary layer and the intensity of the Somali jet. Another is the daily variation of solar radiation, that makes an impact on the pattern and intensity of the airflow over the Somali region. (*b*) The topography is also an important factor in acting on the Somali jet. First, as an obstacle, it forces the eastly airflow flowing over the south of the equator turning to be northward. The thermal contrast between it and the flatlands is favourable to strengthening the Somali jet.

The Interannual Variations of Monsoon and its Physical Causes

The relationship between the interannual variations of the East Asian monsoon and the global atmospheric circulation has been discussed. The variations of the intensity of the East Asian monsoon is linked with the global circulation through the meridional circulations and the teleconnection with ultra-long waves. The variation of heating sources is one of the major causes to induce the interannual variability of monsoon. The activities of the subtropical high over the Western Pacific in summer are related to the coverage of the polar ice cap in winter and the snow coverage over the Tibetan Plateau affects the East Asian summer monsoon circulation to a certain extent. The results from the numerical experiments show that the Asian summer monsoon is weaker in the year of El-Nino than in the other years in general.

The Summer Monsoon and Long-Range Forecast of Rainfall in China

The flood and drought are likely to occur when the monsoon circulation is anomalous. The thermal condition of the underlying surface is one of the primary physical causes to induce the monsoon circulation anomaly. The occurrence of flood and drought in China is closely related to the thermal condition over the Tibetan Plateau, SST over the Kuroshio region, equatorial Pacific and Indian Ocean. For instance, there is a positive correlation between the summer precipitation in Southwest China and the winter SST in the Bay of Bengal. The anomaly of the monsoon circulation is also responsible for the extremely warm and cold summers.

The Onset and the Medium Range Oscillation of the Summer Monsoon

The Asian monsoon system is composed of two sub-systems, which are relatively independent and also interact with each other. The East Asian monsoon, which generally set on earlier than the Indian monsoon, moves westward after onset. The diagnoses of the heat source indicate that the sensitive heat transport over the eastern Tibetan Plateau in April, May, June creates the changes of the atmospheric temperature field, thus results in the occupation of the South Asian high over the plateau. Some numerical simulations further elaborate the onset of the summer monsoon. The sensitive heat transport over the Plateau increases with the intensity increase of the solar incident radiation. The jet is forced to retreat northward, when the sensitive heat transport is strong enough to surpass the adiabatic descending effect in the south branch of the westerly stream jet. That results in the onset of the summer monsoon. The mid-latitude circulation structure gives an apparent intensifying effect on the East Asian monsoon.

During the prevailing season of the summer monsoon, various oscillations with different periods, such as 4-5 days and two weeks, have been found. The periodic variations of precipitation with the same periods in the south of China have also been picked out.

The studies on the interactions between the Indian monsoon and East Asian monsoon suggest that the energy is transported from the area of the South China Sea monsoon to that of the Indian monsoon in the upper troposphere, and it takes a contrary way in the lower troposphere.

The Weather Systems during the Summer Monsoon

The weather systems during the East Asian monsoon are mostly different from that during the Indian monsoon. It is indicated that the rainstorm in South China is related to the activities of the monsoon low and monsoon surge in the South China Sea. The monsoon low in the South China Sea is quite different from that during the Indian monsoon. The release of the latent heat seems to be the principal mechanism of the development and maintainance of the monsoon trough. The analysis of the global cross-equatorial flow shows that the strongest cross-equatorial flows related to monsoon are over Somalia and along 105°E. The intensity of the cross-equatorial flow is basically determined by the pressure patterns in the equatorial area of the south hemisphere. But the cross-equatorial flow over the South China Sea barely affects the intensity of the East Asian monsoon at least in some case studies.

DYNAMICAL PREDICTABILITY OF SEASONAL MONSOON RAINFALL: PROBLEMS AND PROSPECTS

B N GOSWAMI

Centre for Atmospheric Sciences, Indian Institute of Science, Bangalore-560 012, *India*

(*Received* 28 *June* 1993; *Accepted* 07 *July* 1993)

The conceptual basis and the current status of dynamical prediction of seasonal mean monsoon rainfall are reviewed. The monsoon may be viewed as a combination of a planetary scale wave number one circulation and a regional fluctuating Hadley circulation. The planetary scale component is primarily governed by the sea surface temperature in the Pacific and the general circulation models are successful in simulating it. The models, however, have significant systematic errors in simulating the regional component which seems to be sensitive to a number of physical processes. It is noted that the simulation of mean rainfall over Indian region is quite sensitive to initial conditions while those over Pacific or Sahel are not. It is hypothesised that, while over most of the tropics the low frequency variability is governed mainly by slowly varying boundary conditions, the internal dynamics plays an important role in determining even the low frequency variability over the Indian region. Evidences that there are strong intraseasonal oscillations over Indian region during northern hemispheric summer and that these intraseasonal oscillations arise due to internal dynamics are presented. A conceptual picture of the intraseasonal oscillations of the monsoon is presented and it is indicated that several factors of comparable strength may influence the intraseasonal oscillations. Due to these inherent limitations, new techniques such as ensemble forecasting have to be evolved for better prediction.

Key Words: Dynamical Predictability; Seasonal Monsoon Rainfall; Planetary Scale Wave; Hadley Circulation; Low Frequency Varbiability; Intraseasonal Oscillations; Reverse Hadley Cell; Equatorial Circulation

Introduction

Characteristically, the day to day fluctuations of the tropical atmosphere (i.e., the tropical weather) are much smaller than those in the extratropics. For example, large day to day fluctuations in temperature has a large impact on the day to day life, energy requirements, transportation and agriculture in the extratropics. In the tropics, we rarely experience such large temperature fluctuations. The only tropical system that has significant effect on society is the tropical cyclone. Tropical cyclones spend most of its life over oceans and affect some coastal areas and some islands. As a result, tropical weather has a relatively lesser impact on society as compared to middle latitude weather. However, the variations of the mean conditions of the atmosphere (or the changes in the climate) in the region can have a much stronger impact on the society. This is mainly through changes in the amount and distribution of the mean rainfall as it affects the agricultural production and availability of ground water, safe drinking water and shelter. (For an example of how adverse monsoons affect the Indian economy see Swaminathan[1]). This is why the major demand on the tropical meteorologists has been for the long range prediction of the amount and the distribution of the mean rainfall. This demand has been even stronger for the long range prediction of the mean monsoon rainfall as monsoon is the most vigorous and powerful regional circulation in the tropics having enormous societal impact.

Starting with Blandford[2] and followed by Walker[3,4], India has a long tradition of long range prediction of monsoon rainfall. Recognising that various planetary scale phenomena influence the monsoon circulation, statistical correlations between Indian monsoon rainfall (NIR, Normalized all India seasonal mean Rainfall) and a large number of regional and global parameters have been identified and used in multiple regression formulae[5-7] and more recently in power regression formulae[8,9]. In the absence of better tools for long range prediction, these statistical techniques have been used by India Meteorological Department as operational tools for long range prediction of NIR. In fact the Gowarikar *et al.*[9] power regression model has been quite successful in predicting the NIR in May for the last five years (1988-92). However, these statistical models are hardly ever reliable over a long stretch of time. The reason behind

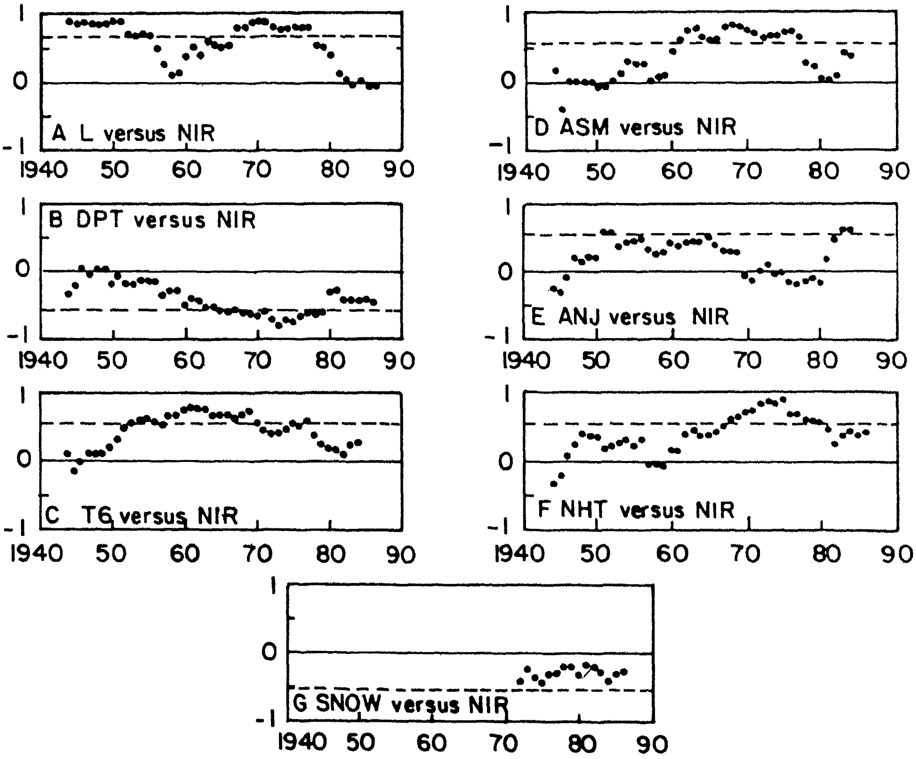


Fig 1 Eleven year sliding correlations between Indian monsoon rainfall NIR and different predictors. Values are plotted at the central year of 11-year intervals. (A) Latitude position of 500mb ridge along 75°E in April, L (B) Pressure tendency (April minus January) at Darwin, DPT (C) March-April-May temperature at six stations in west central India T6 (D) SST anomaly in north eastern Arabian Sea in May, ASM; (E) SST anomaly in Arabian Sea in January ANJ (F) Northern Hemisphere temperature anomaly in January-February NHT (G) Eurasian snow cover in January (after Hastenrath and Greichar[10]).

this is related to the fact that the statistical relationship between almost all predictors with NIR undergo significant low frequency changes[10]. Fig. 1 taken from Hastenrath and Greischar[10] illustrates this point. The correlations often change from significantly positive (negative) to zero or negative (positive). In view of these slow changes in the predictor parameters associated with slow climatic changes, the Gowarikar *et al.*[9] model should also be studied for stability and its performance should be tested for a larger set of independent samples (e.g., using methods similar to one used by Kung and Sharif[11]).

With the recognition of the limitations of the statistical models for long range prediction, the search for an alternative method has been on for quite some time. This has led to the birth of dynamical climate prediction using deterministic models[12-14] in the mid eighties and the recent years have seen considerable growth of the field. While, dynamical (or numerical) weather prediction started way back in late fifties and early sixties and has made tremendous progress over the years, why did dynamical climate prediction make such a late start? We can think of at least two reasons for this.

(1) The primary reason has been that until the early eighties, we had a kind of "conceptual barrier" for deterministic climate prediction. This was because, the deterministic limit on predictability of the atmosphere[15-17] was well established which showed that due to the intrinsic nonlinearity of the atmosphere, the predictability of the instantaneous state of the atmosphere is limited to two to three weeks. Climate prediction requires integration of the dynamical models upto several months. Since the instantaneous state of the atmosphere loses its memory regarding its initial condition, after about 2-3 weeks, it was thought that climate prediction may be impossible. It was Charney and Shukla[18] and Shukla[19] who showed that while the detailed structure of the flow patterns cannot be predicted beyond a couple of weeks, the space time averages of the weather elements may be predictable beyond this limit. This is because of the following.

While the day to day fluctuations of the atmosphere is governed by synoptic hydrodynamical instabilties and their saturation mechanisms, that of the monthly and seasonal means is governed by the low frequency planetary scale flow patterns. The predictability of planetary scale is governed partly by "internal dynamics" or instability of their own scale or interactions with highly unstable synoptic scale and partly due to "slowly varying boundary forcing." The slowly varying boundary forcing such as the sea surface temperature (SST), soil moisture, sea ice, snow cover etc. can produce anomalous sources and sinks of heating in the planetary scales and thereby modulate the evolution of planetary scale flow in a slow time scale. If the growth and decay of the planetary scales were purely governed by internal dynamics, there would not be much hope for their prediction beyond the limit of deterministic predictability. On the other hand, if the planetary scale patterns are governed by slowly varying boundary forcing, it can provide significant coherent changes of the planetary scales at low frequencies leading to potential for predictability of the space time averages.

Charney and Shukla[18] indicated that the monthly or seasonal mean in the middle latitudes may be mainly governed by internal dynamics while those in the tropics may be governed mainly by boundary forcing. This is because, the middle latitude dynamics is governed by large amplitude baroclinic instabilities with time scales of the order of a week and the variations of the boundary conditions such as SST, sea ice and snow cover cannot effectively translate into heating anomaly and then to circulation anomaly. On the other hand, the tropical dynamics is governed by convective instabilities which have much shorter period (of the order of a day or less) and the boundary forcings such as SST, soil moisture etc. can effectively translate themselves into heating anomalies and circulation anomalies due to the higher mean tropical temperatures. Thus, the conceptual foundation for climate prediction was laid. This line of argument leads to the conclusion that in the short and medium range, the tropics may be less predictable than the extratropics due to the short time scale of the convective instability compared to that of the baroclinic instability. On the other hand, the monthly and seasonal means in the tropics may be more predictable than those in the extratropics, mainly due to the dominance of boundary forcing in the tropics. In fact many studies[20-24] that followed the pioneering work of Charney and Shukla[18] have confirmed the above conclusion.

(2) The secondary reason for the late start of climate prediction efforts was the availability of computer technology. For climate prediction complex general circulation models (GCMs) have to be integrated for durations of many months. It demands large dedicated computing facility. In the early eighties, this type of facility was available only to a few select groups.

Over the last decade, computing technology has advanced at a breathtaking speed. However, the progress in the field of climate prediction has not been as fast. This is related to, among others, two problems.

(*a*) The major problem in the progress of climate prediction has been the systematic errors or climate drift of existing GCMs. This means that given observed boundary conditions and observed initial conditions, almost all GCM simulations of the present day climate has significant differences with the observed present day climate[25,26]. This makes it difficult to make predictions for future climate with these models. In some cases, prediction experiments have been carried out by artifically removing the climate drift (e.g., Shukla[27]). Since such procedures do not add to our understanding, continuous efforts are being made to remove these climate drifts by improving the physics of the models. However, the physical processes responsible for the climate drifts in various models are not always transparent. Continuous and systematic experimentation and diagnosis are required to identify the cause behind the climate drift. This is one reason for slow progress in this field.

(*b*) For long range predictions with lead time of a season or more, the boundary conditions for the atmosphere do not remain steady for such a period, i.e., for time scales of a season or longer the atmosphere, ocean and land interact strongly specially in the tropics. Again, the primary reason for this is the large mean SST in the tropics. A change in SST by 1-2°C around a mean of 27°C can result in a large heating anomaly in the tropics whereas a change in SST as large as 4-5°C around a mean SST of 20° will not result in any appreciable heating anomaly in the middle latitude. Therefore, the coupled ocean-land-atmosphere system has to be considered for climate prediction. As a result, the need for understanding the physics of the coupling between different components of the climate system (such as between atmosphere and ocean and between atmosphere and land) and the ability to model them arises. During the past two decades considerable progress has been made in our understanding of the tropical coupled ocean-atmosphere system[28-33] and the land-atmosphere system[34-36]

Having recognised that for long range prediction we have to consider the coupled climate system, the concept of "boundary forced" predictability as introduced by Charney and Shukla[18] has to be modified. In the coupled system, the boundary condition for one component is a field variable for another component. In other words, all the variables are part of a self-consistently evolving system. For example, in the coupled ocean atmosphere system in the tropics, the SST may be considered as a boundary condition for the atmosphere. SST influences the atmospheric heating distribution which determine the surface wind distribution. The surface winds drive ocean currents which eventually changes the SST. The changed SST changes the atmospheric heating and the coupling continues. Therefore, the concept of boundary forced predictability cannot be applied literally. However, it turns out that the coupled system possesses some natural slow modes of oscillation of the system[33-39]. Such natural slow modes of oscillation modulate the space time averages of the system at slow time scales and leads to long range predictability of the system[40].

With the conceptual framework in place, the dynamical climate models becoming better and better and with faster computing facilities being available, all the elements needed to start seasonal prediction of monsoon using dynamical models are available. Therefore, we may expect such dynamical seasonal monsoon prediction to be operationally feasible in some foreseeable future.

Dynamical Prediction of Mean Monsoon: Status

However, the road towards dynamical prediction of mean monsoon has not been smooth. This is because, although most climate models have improved considerably over the past decade and reduced climate drift in other parts of the globe, the simulation of the mean monsoon has been one of the most challenging problems. In the mid eighties, many monsoon meteorologists recognised that there is a need to conduct systematic studies and cooperative efforts to identify the intrinsic problems of simulation of the monsoon circulation and to improve the simulations. This resulted in the birth of the Monsoon Numerical Experimentation Group (MONEG) under the World Climate Programme (WCP) of the World Meteorological Organization. With the initiative taken by MONEG a series of monsoon simulations were carried out by a number of General Circulation Modelling groups around the world to simulate the 1987 and 1988 monsoons using identical boundary conditions and identical initial conditions. The results were compared in a workshop[41]. The diversity in simulation of the monsoon circulation by different GCMs is illustrated in Fig. 2 where the strength and position of the Somali Jet and Tropical Easterly Jet (TEJ) simulated by different models are shown. This shows that there are significant differences in simulating the mean monsoon by different models. Not only that, even two different analyses such as NMC and ECMWF sometimes do not agree with each other. This indicates that there is a need to reanalyse past observations over a period of time using a standard analysis technique. The main success of the models has been in—

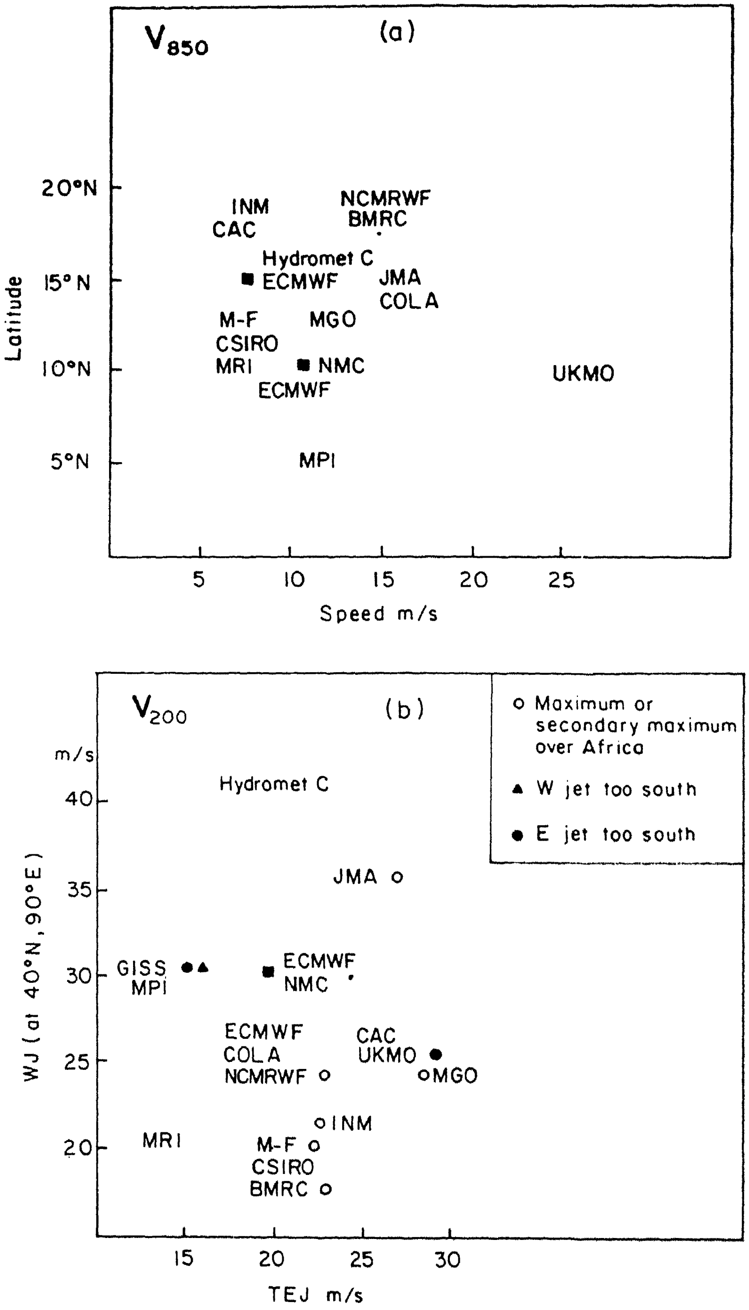


Fig 2 Simulated and analysed wind over India for June-July-August 1988 (a) at 850 hPa and (b) at 200 hPa. Abscissa is maximum wind speed of Somali Jet (a) and Tropical Easterly Jet, TEJ (b). The ordinate is latitude of the maximum wind (a) and strength of westerly jet (b). Squares represent analyses from ECMWF and NMC. The acronyms are explained in *Appendix* and represent general circulation models of these groups. (From WMO[41].)

(*a*) Simulation of the planetary scale features of the monsoon as evidenced by the 200 hPa velocity potential[41,24].

(*b*) Majority of the models also could simulate the correct tendency (not magnitude) of the interannual variability between 1987 and 1988. (For example, they were able to simulate stronger easterly jet and more intensified monsoon flow into Africa and wetter conditions during 1988 as compared to 1987[24,41]. However, most models still have significant systematic errors in simulating the regional feature of the monsoon.)

(*c*) Although some models are now beginning to get the two centres of precipitation one in the western side of Western Ghat and the other over the head of Bay of Bengal (e.g., Palmer *et al.*[24]), most models have a dry bias over the continent[24,41].

(*d*) Most models either fail to simulate the quasi-stationary monsoon trough or simulate it too far to the north. The dry bias over the continent may be related to the models' inability to simulate the monsoon trough.

(*e*) As shown in Fig. 2, the models also have some systematic errors in simulating the strength and position of the Somali Jet and the TEJ.

The systematic errors associated with the regional scale circulation are found to be sensitive to parameterization of the physical processes such as convection, surface heat fluxes over the warm pool (Indian Ocean and Western Pacific) region, land surface processes and representation of subgrid scale orography. Another rather interesting lesson learnt during the exercise is that,

(*f*) The planetary scale circulation (e.g., 200 hPa velocity potential is not sensitive to small changes in the initial condition, but the regional mean monsoon circulation (e.g., rainfall) is quite sensitive to such changes in the initial condition. However, the mean rainfall over Sahel is not so sensitive to initial conditions[24,42]. This is illustrated in Fig. 3 (taken from Palmer *et al.*[24]). This figure shows the differences in the mean July + August simulated rainfall for Indian and African region between two simulations which were identical except

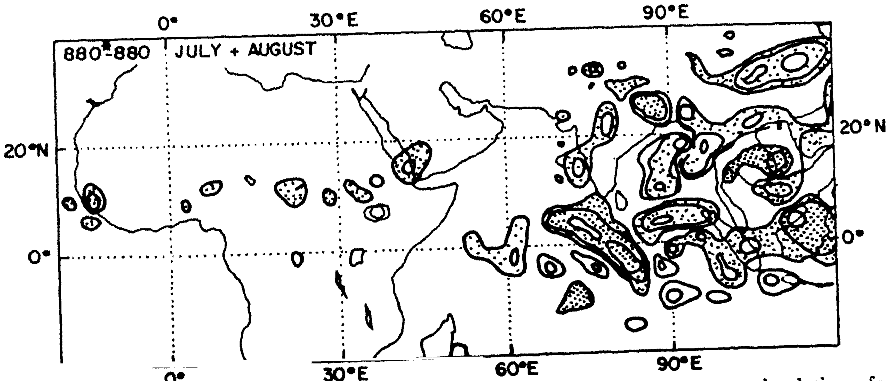


Fig 3 Rainfall differences averaged between simulated days 31-90 between two simulations for July and August of 1988 with observed sea surface temperatures as boundary conditions but with initial conditions corresponding to June 1, 1988 (880*) and June 2, 1988 (880) respectively. Coarse stipling for values greater than 2mm day^{-1} and fine stipling for value less than -2mm day^{-1}. Contours at 2, 4, 8, 16mm day^{-1}. (after Palmer *et al.*[24]).

that their initial conditions were separated by one day. It is seen that there is a large difference over the Indian region while there is hardly any difference over the Sahel region.

The last point is quite intriguing and may be indicative of reasons behind difficulties in simulating the mean monsoon circulation. The low frequency variations over most of the tropics (e.g., Pacific and Sahel region), are mainly governed by slowly varying boundary conditions rather than internal dynamics[24,42-43]. Over the monsoon region, however, the internal dynamics seems to play an important role. This realization is based on the observation that during the NH summer, the Indian monsoon region is the region strongest in intraseasonal oscillation (ISO) activity. There exists considerable amount of theoretical and observational evidence that most of the ISOs in the tropics are generated due to internal dynamics of the system.

To understand the problems of simulation of the mean monsoon circulation, therefore, we need to understand the nature of the ISO of the monsoon and the physics that govern them. In the next two sections, we summarise the observed characteristics of the ISO and the theoretical and conceptual models available so far to understand them.

Intraseasonal Oscillations of Monsoons: Summary of Observations

The fluctuations of the monsoon manifest in the form of active and weak (or break) spells of monsoon rainfall (Fig. 4).

(*i*) These active/break spells of monsoon are associated with fluctuations of the regional tropical convergence zone (TCZ) or local Hadley circulation[44-49].

(*ii*) There appears to be two favourable locations of the TCZ, one over the continent and another over the equatorial Indian Ocean[44,47,48]. This is also supported by an anti-correlation of daily OLR over the continent and the equatorial Indian Ocean[44,49].

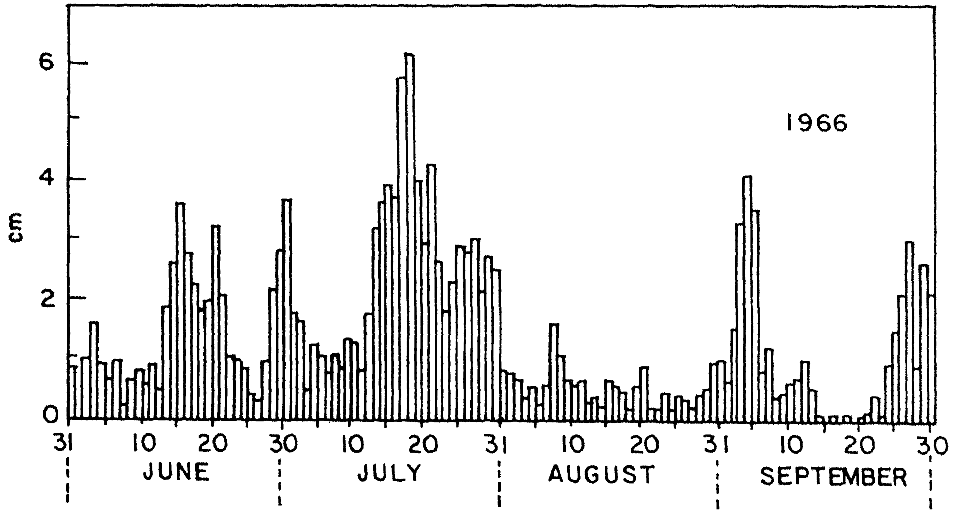


Fig 4 Time series of daily rainfall over northern India showing active/break spells in the monsoon rainfall for 1966.

(*iii*) The fluctuations of the TCZ over the monsoonal area is due to *in situ* generation and demise of the continental TCZ as well as repeated northward propagation of the oceanic TCZ from the equatorial Indian Ocean[44-47].

(*iv*) The ISO of the monsoon, though not periodic, has two dominant bands in the spectrum[45,50-53]. One band contains periods between 10-20 days while the other contains periods between 30-50 days. The 10-20 day oscillation does not seem to have significant northward propagation while it has a westward propagation. The 30-50 day oscillation has a dominant northward and eastward propagation.

The maximum cloud zone or the TCZ represents a region of large scale organized convection and precipitation. The latent heat release associated with the precipitation also makes this the location of organized large scale deep heating of the atmosphere. Thus, the maximum cloudiness (or the TCZ or the heat source) averaged over the whole season though appears to be located over continental India, in reality is fluctuating between the two preferred positions, the continental one and the oceanic one. It turns out that the continental heat source dominates over the oceanic heat source for a longer period of time within the season. As a result, averaged over the whole season, the heat source appears to be on the land. During an active spell, the continental heat source dominates while during a break spell the oceanic one dominates.

Before we could understand the origin of the ISO we must have clear picture of the mean monsoon circulation. In the next section, we present a conceptual picture of the mean monsoon circulation.

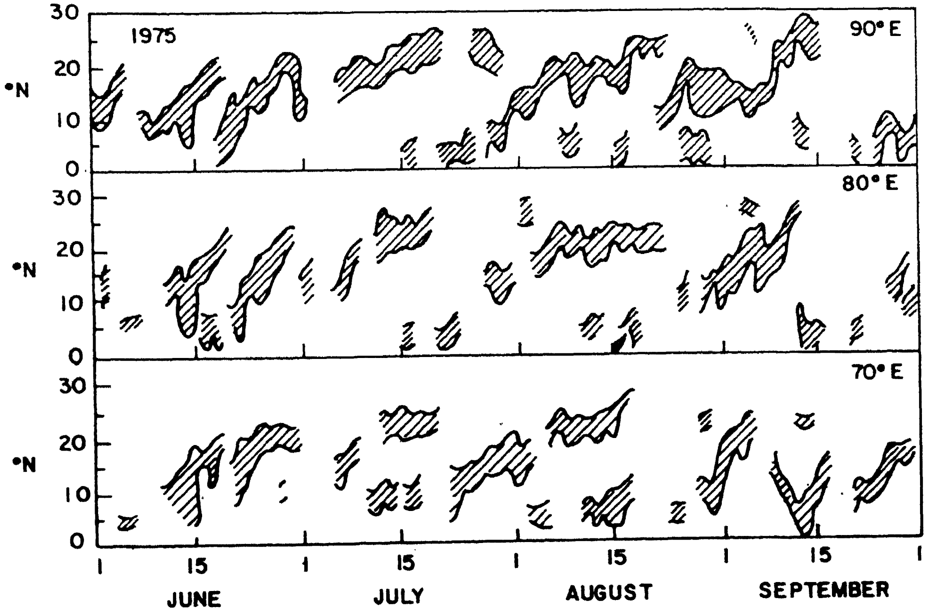


Fig 5 Evolution of the maximum cloud zone in 1975 along three Indian longitudes showing repeated northward migration as well as *in situ* generation (after Gadgil[48]).

The Mean Monsoon

Fig. 6 (taken from Murakami[54]) shows the climatology of the seasonal march of the outgoing long wave radiation (OLR). It shows how the region of organized deep convection (OLR values less than $240W/m^2$) moves from about 5°S in December-January-February to about 25°N in June-July-August over the Indian/Indonesian region. This is associated with an identical movement of the vertically averaged mass-weighted diabatic heating rate (Fig. 7, from Johnson *et al.*[55]). The monsoon is essentially a direct circulation driven by these heat sources. The important point to be made is that the heat source is nearly zonal in winter situated at about 5°S extending from about 30°E to about 140°W. This heat source will produce a Hadley circulation in the north-south and vertical plane with ascending motion around 5°S and descending motion at about 30°N over this whole region. Since this heat source is closer to the equator, it would also produce an equatorial east west response resulting in a Walker circulation with ascending motion over this region and descending motion in the eastern Pacific. This would merge with the descending motion resulting from the Walker response to the South American heat source. In contrast, during summer the heat source is oriented from south east centered at the equator and 160°E to north west centered at about 25°N and 75°E. In other words, the mean position of the tropical convergence zone (TCZ) is oriented from south east to north west direction in summer. As a result, the mean Hadley circulation associated with the TCZ has changed drastically over this region in summer. While

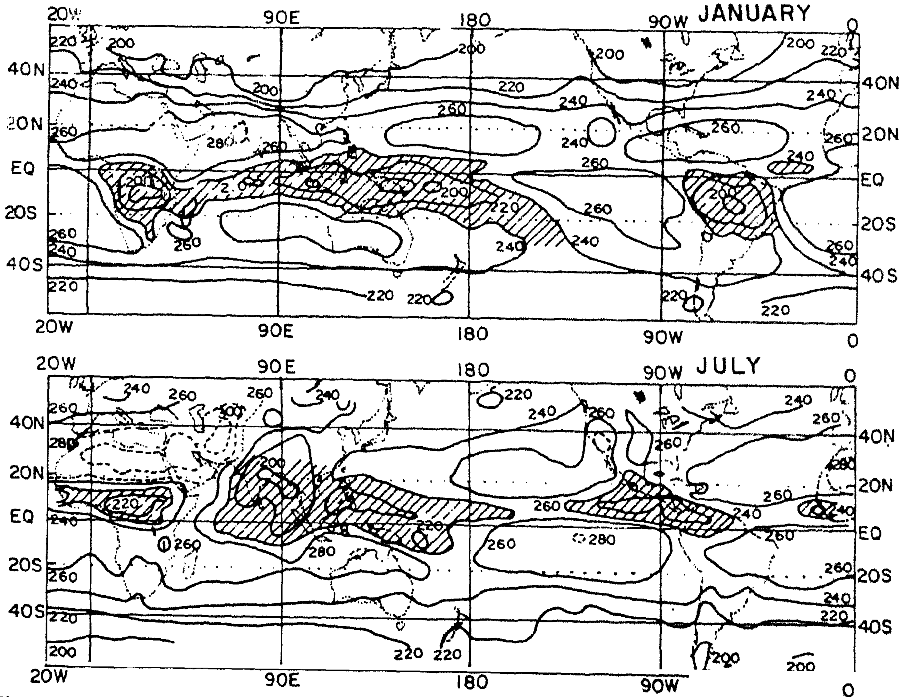


Fig 6 Mean outgoing longwave radiation (OLR) for (A) January and (B) July. Tropical regions with OLR less than $240Wm^{-2}$ are shaded and represent regions of deep convection.

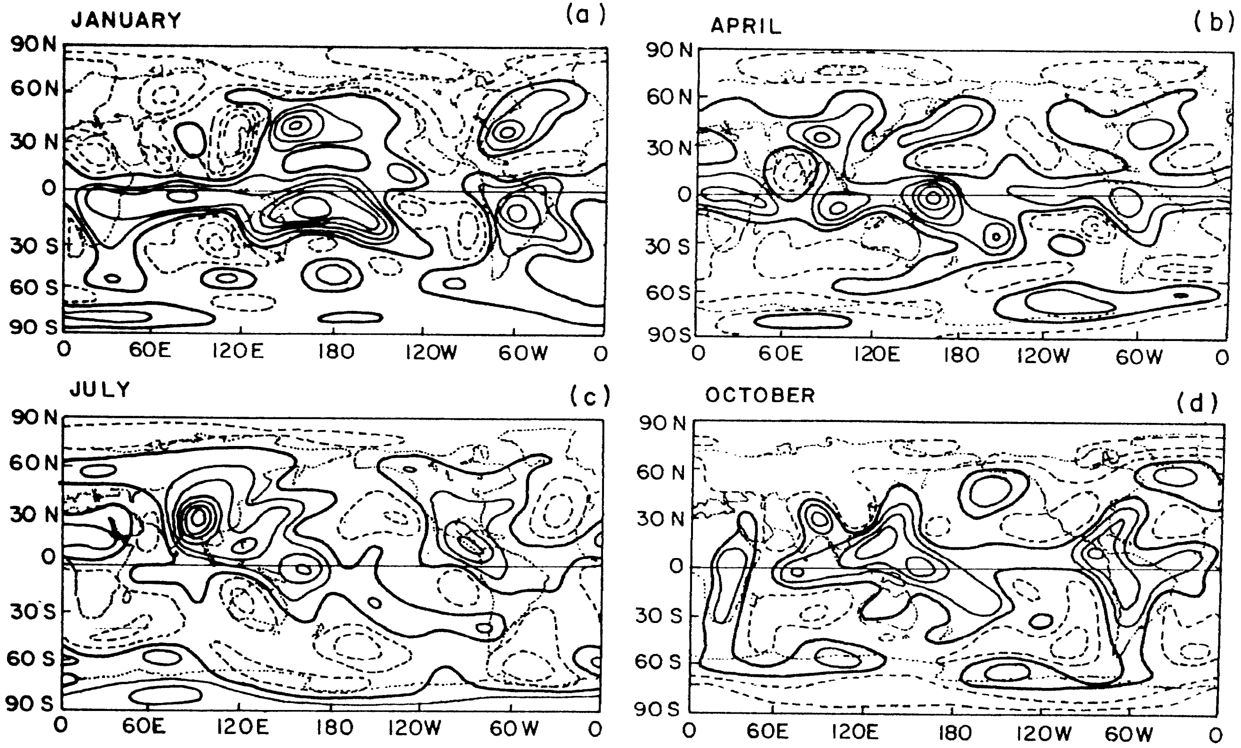


Fig 7 Mass weighed, vertically averaged diabatic heating rates (K per day) computed from ECMWF Level III analyses for four months of 1979. The vertical mass flux is set to zero at the upper boundary of 400 K. The data were filtered to emphasize heating contributions from wave lengths greater than 400km. The zero contour is thick solid line, negative contours are dashed lines and positive contours are thin solid line. Contour interval is 0.5°K day^{-1} (From Johnson *et al.*).

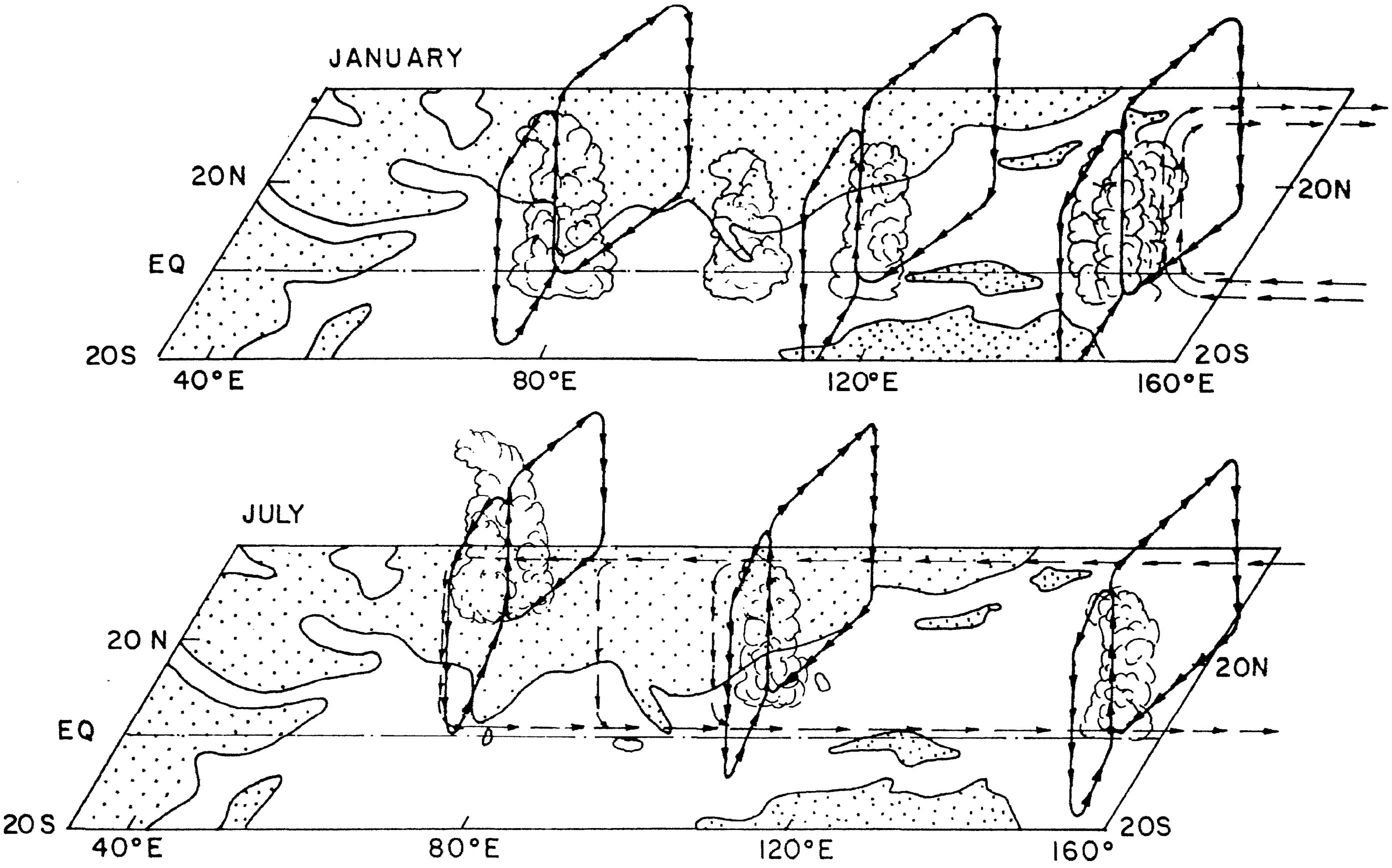


Fig 8 A schematic representation of the mean monsoon circulation during winter (top panel) and summer (bottom panel). It can be seen as interplay between a north south Hadley circulation and an east-west Walker circulation.

there is a regular Hadley cell in the western Pacific with ascending motion around equator and descending motion around 25-30°N, over the Indian region we now have the so called 'reverse Hadley cell' with ascending motion around 25°N and descending motion near equator. Moreover, we note that during winter, the ascending branch of the Walker cell extends from western Pacific to equatorial central Indian Ocean. However,, during NH summer, the heat source situated at about 25°N can have zonal response confined mainly to the western side of the heat source. This is due to the fact that Kelvin waves that are responsible for setting up the response to the east of the heat source are not being allowed at such latitudes. Therefore, the ascending limb of the equatorial Walker circulation in summer is confined to the western Pacific and would have a descending limb over the central Indian Ocean and over eastern Pacific. Over the equatorial Indian Ocean, the descending branch of the local Hadley circulation coincides with the descending branch of the Walker circulation. This is schematically shown in Fig. 8. With this mean picture in mind, let us now try to understand fluctuations of the monsoon.

A Conceptual Framework for Low Frequency Monsoon Fluctuations

With this background, we can now construct a conceptual model of the low frequency variability of the monsoon (Fig. 9). We know that monsoon is a direct circulation forced by deep tropospheric heat sources. Thus, the low frequency variability of the monsoon must be due to the low frequency variability of the heat sources. As we have noted above, there are two distinct large scale regional heat sources, the continental heat source associated with the continental preferred position of the TCZ and an oceanic heat source associated with oceanic preferred location of the TCZ. The next question then is: what causes the low frequency fluctuations of these heat sources? Although radiation plays a role in selecting the location of these heat sources, these deep heat sources are basically maintained by organized convection. Organized convection is basically driven by net moisture convergence into the region. Thus, low frequency variability of the heat sources are in turn caused by low frequency variability of the moisture convergence. What are then the factors that influence the low frequency variability of the moisture convergence?

The moisture convergence for the oceanic heat source is governed by the following:-

(*i*) Large scale convergence in the near equatorial region.

(*ii*) Cloud-radiation feedback.

(*iii*) Evaporation from regional SST anomalies.

(*iv*) Interaction between dynamics and organized convection.

The large scale convergence in the equatorial region is governed by—

(*a*) Subsidence from the continental heat source. In this way the continental heat source and the oceanic heat source interact with each other.

(*b*) By the global 30-50 day eastward propagating oscillation. The wavenumber one convergence/divergence anomalies associated with this mode passes through the equatorial Indian Ocean once in 30-50 days. In the convergent phase, it would try to favour the equatorial heat source while in the divergent phase it would try to suppress it.

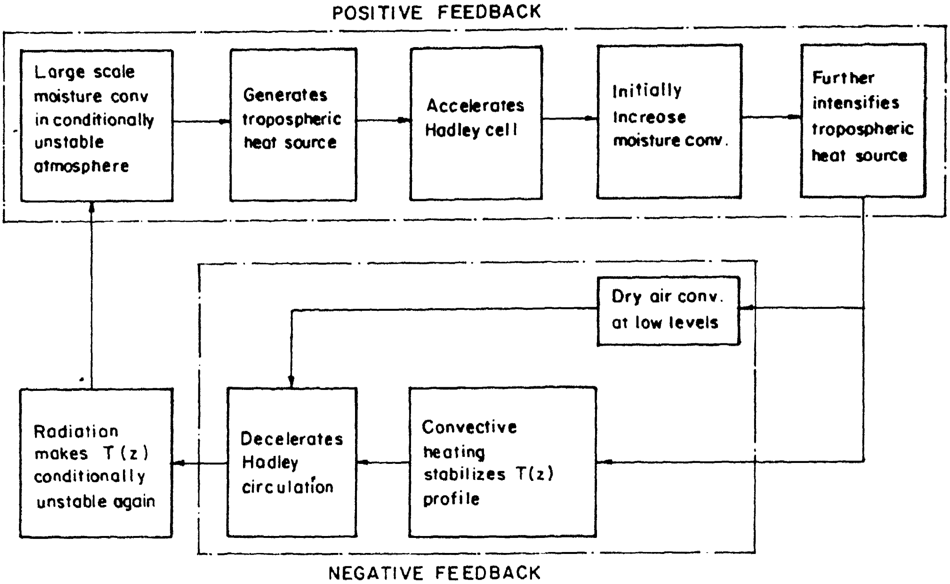

POSITIVE FEEDBACK
Large scale moisture conv in conditionally unstable atmosphere
Generates tropospheric heat source
Accelerates Hadley cell
Initially increase moisture conv.
Further intensifies tropospheric heat source
Dry air conv. at low levels
Radiation makes T(z) conditionally unstable again
Decelerates Hadley circulation
Convective heating stabilizes T(z) profile
NEGATIVE FEEDBACK

Fig 9 A schematic representation of the proposed feedback mechanism for *in situ* oscillation of the Hadley circulation or the TCZ due to organized convection and large scale dynamics (Goswami and Shukla[57]). This mechanism can operate over the ocean as well as over land.

(*c*) *Walker circulation*: During summer in a non-El Nino year, the equatorial Indian Ocean is a region of subsidence associated with the Walker circulation with its ascending branch over western Pacific. In an El Nino year, the maximum precipitation zone moves to eastern Pacific zone, the ascending branch of the Walker circulation also moves to the eastern Pacific. Hence, the equatorial Indian Ocean receives much less large scale subsidence. With the reduced subsidence, the high SST over this region tends to favour the oceanic heat source and produce the oceanic TCZ. This may lead to dominance of the oceanic heat source over the continental one leading to a poor monsoon over India during a warm episode of the ENSO cycle.

We note that the interannual variations of SST over this region is of the order of 0.5°C or less (e.g., Rao and Goswami[56]). Thus, evaporation anomalies are not very large and the large scale convergence anomalies may be of comparable amplitude or larger.

As shown by Goswami and Shukla[51], interaction between dynamics and organized convection itself can result in low frequency fluctuations of the heat source through fluctuation of the moisture convergence. We shall discuss this mechanism in some detail in the next section.

The moisture convergence for the continental heat source is governed by the following:-

(*a*) Large scale convergence in the monsoon trough region.
(*b*) Cloud radiation feedback.
(*c*) Evaporation associated with land surface processes.
(*d*) Interaction between dynamics and organized convection.

The moisture convergence over the monsoon trough region is modulated by:

(i) subsidence from the oceanic heat source. This is the mechanism through which the oceanic and the continental heat sources interact amongst each other; and

(ii) through interaction with middle latitude circulation. When the middle latitude waves penetrate to this latitude, they can influence the low level convergence in this region. Recently, Sankar Rao *et al.*[58] showed that appreciable influence of the middle latitude waves reach monsoon region during break spells.

The same mechanism, mentioned above, that effects the oceanic heat source through interaction between dynamics and organized convection can also produce low frequency fluctuations of the continental heat source.

Thus, the fluctuations and interactions between these two large scale regional heat sources are governed by several complex interactions. Some of them are regional in nature (e.g., local evaporation, ground hydrology, cloud-radiation feedback etc.) while some of these interactions are influenced by planetary scale processes (e.g., the global 30-50 day effect, the role of the equatorial Walker circulation and the middle latitude influences.)

The conceptual picture presented above describes the dynamical factors that may make the continental and the oceanic heat sources fluctuate *in situ*. This picture is further complicated by the northward propagation of the ocea-

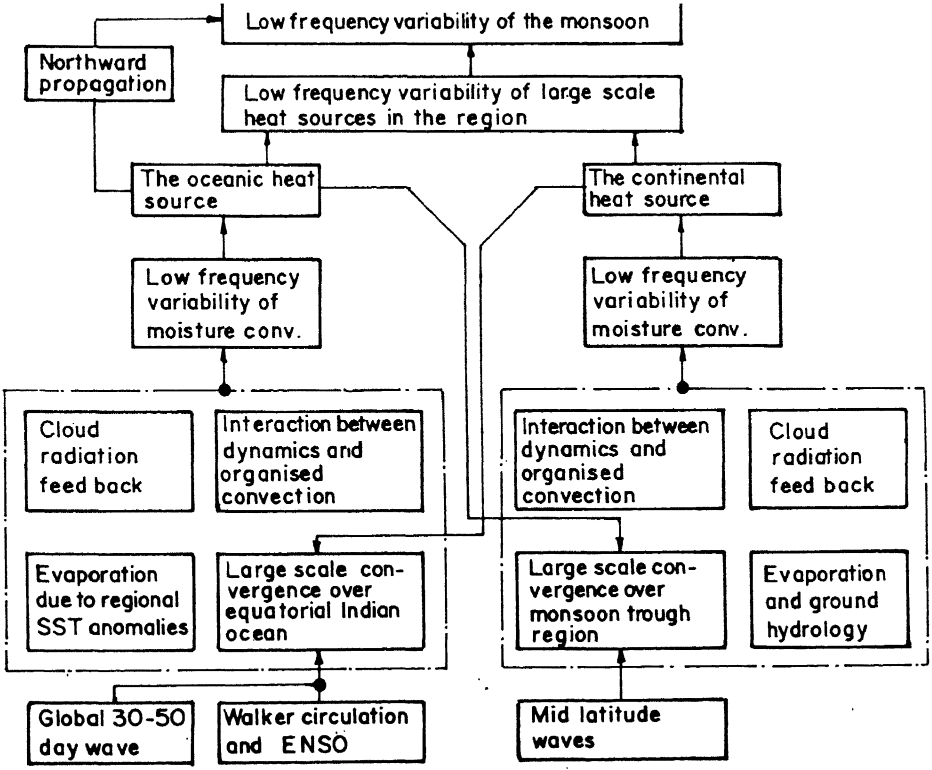

Low frequency variability of the monsoon
Northward propagation
Low frequency variability of large scale heat sources in the region
The oceanic heat source
The continental heat source
Low frequency variability of moisture conv.
Low frequency variability of moisture conv.
Cloud radiation feed back
Interaction between dynamics and organised convection
Interaction between dynamics and organised convection
Cloud radiation feed back
Evaporation due to regional SST anomalies
Large scale convergence over equatorial Indian ocean
Large scale convergence over monsoon trough region
Evaporation and ground hydrology
Global 30-50 day wave
Walker circulation and ENSO
Mid latitude waves

Fig 10 A schematic diagram showing factors responsible for low frequency variations of the monsoon and interactions between them.

nic TCZ (or the heat source). Even if there was no *in situ* fluctuations of the continental heat source, repeated northward propagation from the ocean to the land of the oceanic TCZ would itself generate certain low frequency variability over the monsoon region.

Mechanisms of ISO of Monsoon: Role of Internal Dynamics

In order to understand the ISO of the monsoon we must explain

(*i*) why there are two preferred positions of the heat source or the TCZ and what maintains them;

(*ii*) what are the mechanisms that cause *in situ* fluctuations of the individual heat sources; and

(*iii*) what causes the northward propagation of the oceanic TCZ

Existence of the Oceanic and Continental Heat Sources

We recall that during the monsoon season, the maximum SST (about 28°C) in the north Indian Ocean resides just south of India between about 70°E and equator to about 110°E and 15°N. Various studies of the dynamics of the mean TCZ[59] show that in an ocean covered earth, the mean TCZ occurs at the location of the maximum SST. Thus the ocean south of India is a favourable location for the formation of the TCZ. However, the special geography of the Indian continent with the Himalayas to the north and east favours another heat source. Due to the existence of the Himalayan range to the north and east of India, a large scale low level convergence zone known as the monsoon trough is created. This provides the seed for the continental heat source. It provides the small amount of low level lifting necessary for the conditional instability to be effective in producing organized convection. The land surface heating alone produces only a shallow circulation. The monsoon trough helps to convert this shallow circulation into a deep circulation by producing a deep heat source.

Mechanisms of in situ *Fluctuation*

Krishnamurti and Bhalme[50] proposed a mechanism based on cloud radiation feedback to explain the quasi biweekly oscillation. Recently, Goswami and Goswami[60] have proposed a mechanism for the quasi biweekly oscillation as an internal mode of oscillation of the system. Goswami and Shukla[57], Gadgil and Srinivasan[61], Nanjundaiah *et al.*[62], Keshavamurty *et al.*[63], through different modelling studies show that the 30-50 day oscillation of the monsoon can arise due to interactions between organized convection and large scale dynamics. Detailed numerical experimentation with a zonally symmetric GCM in Goswami and Shukla[57] has led to a plausible scenario by which the tropical Hadley circulation or the TCZ could oscillate *in situ* through interaction between organized convection and dynamics. This is schematically shown in Fig. 9. All these studies indicate that these oscillations of the monsoon represent some internal modes of the system.

Northward Propagation

Webster[65] and Goswami and Shukla[57] showed that ground hydrology is crucial for northward propagation of the TCZ. Further insight was provided by Gadgil and Srinivasan[61] where they showed that the meridional gradient of the convective stability factor is important in the northward propagation.

Based on knowledge gained from all these studies we can conclude that fluctuations of monsoon are linked with fluctuations of the TCZs or associated heat sources in this region. As we noted, there are two such heat sources, a continental heat source and an oceanic heat source. As shown by these studies and as discussed in the previous section, the fluctuations of the heat sources are governed by several convective and radiative feedback processes. These feedbacks are sensitive functions of the dynamical flow fields in the region. Thus, the ISO of monsoon is primarily governed by internal dynamics of the atmosphere.

Conclusions

In the absence of any better tool, statistical long range prediction of seasonal monsoon rainfall has been done for more than a century. Limitations of the statistical techniques are well recognised and an alternative method based on the physical laws of the land-ocean-atmosphere system is sought. A well-founded conceptual basis for such dynamical long range prediction in the tropics is in place and mathematical tools and computing facilities are being available to make such predictions. The dynamical long range prediction at the present time is where the numerical weather prediction (NWP) used to be in the early sixties. We expect rapid progress in this field in the next few years. Extensive modelling work in the last few years to simulate and predict mean monsoon circulation have, however, come up with a few surprises. We have learnt that the monsoon is not only the strongest of the regional circulations in the tropics, but also probably the most unique. While the mean circulation in most of the tropics (such as the Pacific and Sahel region) is governed mainly by boundary conditions and are not sensitive to initial conditions, the simulation of mean monsoon circulation seems to depend sensitively to changes in initial conditions.

1. Since we are primarily concerned with the prediction of seasonal mean rainfall, we are mainly interested in simulation of the interannual variation of the monsoon. However, monsoon has strong ISOs with periods between 10-20 days and 30-50 days. Since these oscillations are supersynoptic oscillations, they do not get averaged out in the monthly or even in seasonal means. In other words, the monthly or seasonal means are strongly affected by these ISOs. Thus, even though part of the interannual variation of the monsoon may be governed by slowly varying boundary forcing (such as the global SST etc.), part of it is strongly linked with the ISOs. Therefore, the GCMs will be able to simulate mean monsoon circulation successfully only if they could also simulate the ISOs. There is evidence for this in recent COLA model results (Fennessy *et al.*[64]) where improvement of monsoon rainfall is related to improvement in the simulation of ISO in the model.

2. Since the ISOs are mainly governed by internal dynamics, they will remain to be sensitive to initial conditions (IC). This would result in sensitive dependence of even the time mean on IC. Therefore, even when GCMs are capable of simulating the ISOs, there would be an intrinsic uncertainty in the simulation of mean monsoon circulation/rainfall.

3. The dry bias over the continent in most models is related to a too weak continental heat source. As we noted, the quasistationary planetary scale low level vorticity associated with the monsoon trough is important in supporting this heat source. Proper representation of orography seems to be crucial for simulation of the monsoon trough. Again there is evidence for this in the recent COLA results (Fennessy *et al.*[64]) where the mean orography seem to have resulted in a slightly better representation of the trough leading to better rainfall over continent.

4. Part of the monsoon circulation appears to be governed by slowly varying boundary conditions and hence predictable while part of it appears to be governed by internal dynamics and hence less predictable. In order to average out part of the unpredictable component and to bring out the predictable component, ensemble forecasting technique may be employed where results of several simulations with different initial conditions but with same boundary conditions are averaged to make the prediction. There is already some indication (Palmer 1993, *personal communication*) that this technique yields much better mean monsoon prediction.

5. There is some indication that instead of starting the forecast, for example from May 1, if it is started several months earlier (say, December 1 or January 1) the simulation of the mean monsoon for June-July-August is better. Whether such long spin-up reduces the dependence of the simulation of mean monsoon circulation on initial conditions should be investigated. This is an area where further work is needed.

Acknowledgement

This work is partially supported by the Department of Science and Technology through a grant. The author is grateful to Professor S Gadgil for some comments on the manuscript. He thanks Ms K Nagarathna for help in preparing the manuscript and N H Saji for critically reading it.

Appendix

The abbreviations in Fig. 1 represent general circulation modelling groups that took part in a coordinated comparison of monsoon simulation. They are

1. **BMRC:** Bureau of Meteorological Research Centre, Melbourne, Australia.
2. **COLA:** Centre for Ocean-Land-Atmosphere Interactions, University of Maryland, USA.
3. **CSIRO:** CSIRO Division of Atmospheric Research, Aspendale, Victoria, Australia..
4. **ECMWF:** European Centre for Medium range Weather Forecasts, Reading, UK.
5. **GISS:** Goddard Institute of Space Sciences, New York, USA.

6. **INM:** Institute of Numerical Mathematics, Moscow, Russia.
7. **JMA:** Japanese Meteorological Agency, Tokyo, Japan.
8. **LMD:** Laboratoire de Meteorologie Dynamique du CNRS. Paris, France.
9. **M–F:** Meteo France, Centre National de Recherches Meteorologiques, Toulouse, France.
10. **MGO:** Main Geophysical Observatory, St. Petersburg, Russia.
11. **MPI:** Max Plank Institute for Meteorologie, Hamburg, Germany.
12. **MRI:** Meteorological Research Institute, Tsukuba, Japan.
13. **NCMRWF:** National Centre for Medium Range Weather Forecasting, New Delhi, India.
14. **NMC:** National Meteorological Centre, Washington DC, USA.
15. **UKMO:** United Kingdom Meteorological Office, Brecknell, UK.
16. **Hydromet C:** USSR Hydrometeorological Centre.

References

1 M S Swaminathan In: *Monsoons* (*Eds.* J Fein and P Stephens) John Wiley (1987) 121-139
2 H F Blanford *Proc R Soc London* **37** (1884) 3-22
3 G T Walker *Mem India Met Dept* **24** (1923) 75-131
4 G T Walker *Mem India Met Dept* **24** (1924) 275-332
5 P Jagannathan *Seasonal Forecasting in India: A Review. Sp Publ* No. DG082/650 IMD (1960)
6 P K Das In: *Monsoons* (*Eds.* J Fein and P Stephens) John Wiley (1987) 549-578
7 J Shukla In: *Monsoons* (*Eds.* J Fein and P Stephens) John Wiley (1987) 52-54.
8 V Thapliyal *Mausam* **41** (1990) 339-346
9 V Gowarikar, V Thapliyal, S M Kulshreshta, G S Mandal, N Sen Roy and D R Sikka *Mausam* **42** (1991) 125-130
10 S Hastenrath and L Greischar *Proc Indian Acad Sci* (*Earth Planet Sci*) **102** (1993) 35-48
11 E C Kung and T A Sharif *J Met Soc Japan* **60** (1982) 672-681
12 K Miyakoda, C T Gordon, R Caverly, W P Stern, J Sirutis and W Burke *Mon Wea Rev* **111** (1983) 846-869
13 K Miyakoda, J Sirutis and J Ploshay *Mon Wea Rev* **114** (1986) 2363-2401
14 M A Cane, S E Zebiak and S C Dolan *Nature* **321** (1986) 827-832
15 E N Lorenz *Tellus* **17** (1965) 321-333
16 E N Lorenz *Tellus* **34** (1982) 505-513
17 J Shukla *Adv Geophys* **28B** (1985) 87-121
18 J G Charney and J Shukla In: *Monsoon Dynamics* (*Eds.*) J Lighthill and R Pearce) Cambridge Univ Press (1981) 99-109
19 J Shukla *J atmos Sci* **38** (1981) 2547-2572
20 J Shukla In: *Problems and Prospects in Long and Medium Range Weather Forecasting* (*Eds.* D M Burridge and E Kallen) Springer Verlag (1984) 155-206
21 M Fennessy, L Marx and J Shukla *Mon Wea Rev* **113** (1985) 858-864
22 J Shukla and M Fennessy *J atmos Sci* **45** (1988) 9-28
23 C K Folland, T N Palmer, D E Parker *Nature* **330** (1986) 602-607
24 T N Palmer, C Brankovic, P Viterbo and J Miller *J Climate* **5** (1992) 399-417
25 G J Boer, K Arpe, M Blackburn, M Deque, W L Gates, T L Hart, H Le Trent, E Roeckner, D A Sheinin, I Simmons, R N B Smith, T Takioka, R J Witherald and D Williamson *J geophys Res.* **97D** (1992) 12771-12786
26 T N Palmer, J Shutts and R Swinbank *QJR Meteorol Soc* **116** (1986) 1001-1037
27 J Shukla In: *Predictability in the Medium and Extended Range Proc Workshop European Centre for Medium Range Weather Forecasting* (1988) 251-261
28 S G H Philander *El Nino, La Nina and Southern Oscillation* Academic Press (1990) 289 pp.

29 J D Neelin, M Latif, M A F Allaart, M A Cane, U Cubasch, W L Gates, P R Gent, M Ghil, C Gordon, N C Lau, C R Mechoso, G A Meehl, J M Oberhuber, S G H Philander, P S Schopf, K R Sperber, A Sterl, T Takioka, J Tribia and S E Zebiak *Climate Dynamics* 7 (1992) 73-104

30 N C Lau, S G H Philander and M J Nath *J Climate* **5** (1992) 284-307

31 S G H Philander, R C Pacanowski, N C Lau and M J Nath *J Climate* **5** (1992) 308-329

32 T Nagai, T Takioka, M Endoh, Y Kitamura *J Climate* 5 (1992) 1202-1233

33 M Latif, A Sterl, E Maier-Reimer and M M Junge *J Climate* **6** (1993) (*In press*)

34 J Shukla and Y Mintz *Science* **215** (1982) 1498-1501

35 R E Dickinson and A Henderson Sellers *Q J R Meteorol Soc* **114** (1988) 439-462

36 C A Nobre, P J Sellers and J Shukla *J Climate* **4** (1991) 957-988

37 A C Hirst *J atmos Sci* **45** (1986) 606-630

38 B N Goswami and S Selvarajan *Geophys Res Lett* **18** (1991) 991-994

39 M J Suarez and P S Schopf *J atmos Sci* **46** (1988) 3283-3287

40 B N Goswami and J Shukla *J Climate* **4** (1991) 3-22

41 W M O *Simulation of Interannual and Intraseasonal Monsoon Variability* WMO/TD No 470 (1992)

42 D P Rowell, C K Falland, K Maskell, J A Owen and M N Ward *Geophys Res Lett* **19** (1992) 905-908

43 N C Lau *Mon Wea Rev* **113** (1985) 1970-1996

44 T Yasunari *J Meteor Soc Japan* **57** (1979) 227-262

45 T Yasunari *J Meteor Soc Japan* **58** (1980) 225-229

46 T Yasunari *J Meteor Soc Japan* **59** (1981) 336-354

47 D R Sikka and S Gadgil *Mon Wea Rev* **108** (1980) 1840-1853

48 S Gadgil *Aust Met Mag* **36** (1988) 193-204

49 S Gadgil and G Asha *J Met Soc Japan* **70** (1992) 517-527

50 T N Krishnamurti and H N Bhalme *J atmos Sci* **33** (1976) 1937-1954

51 M Murakami *J met Soc Japan* **54** (1976) 175-181

52 T N Krishnamurti and D Subramaniam *J atmos Sci* **39** (1982) 2088-2095

53 T N Krishnamurti and R Ardunay *Tellus* **32** (1980) 15-26

54 M Murakami In: *Monsoon Meteorology* (*Eds.* C P Chang and T N Krishnamurti) Oxford Univ Press (1987) 354-401

55 D R Johnson, M Yanai and T K Schaack In: *Monsoon Meteorology* (*Eds.* C P Chang and T N Krishnamurti) Oxford Univ Press (1987) 271-297

56 K G Rao and B N Goswami *Mon Wea Rev* **116** (1988) 558-568

57 B N Goswami and J Shukla *J atmos Sci* **41** (1984) 20-37

58 M Sankar Rao, V N Lykossov, E M Volodin, A E Nikitin, A I Degtiarev and Kusuma G Rao *Adv atmos Sci* **2** (1991) 137-148

59 B N Goswami, J Shukla, E K Schneider and Y C Sud *J atmos Sci* **41** (1984) 5-19

60 B N Goswami and P Goswami *Geophys Res Lett* **19** (1992) 681-684

61 S Gadgil and J Srinivasan *Meteorol atmos Phys* **44** (1991) 119-132

62 R S Nanjundiah, J Srinivasan and S Gadgil *J Met Soc Japan* **70** (1992) 529-550

63 R N Keshavamurty, S V Kasture and V Krishnakumar *Beitr Phys Atmos* **59** (1986) 443-454

64 M J Fennessy, J L Kinter III, B Kartman, L Marx, S Nigam, E Schneider, J Shukla, D Straus, A Vernekar, Y Xue and J Zhou *The Simulated Indian Monsoon: A GCM Sensitivity Study* (1992) *Preprint*, COLA, Dept of Meteorology Univ Md College Park Md 20742 USA

65 P J Webster *J atmos Sci* **40** (1983) 2110-2124

EVOLUTION OF THE BOUNDARY LAYER OVER THE SOUTHERN OCEAN

M CHATTERJEE

Centre for Atmospheric Sciences, University of Calcutta 92, A.P.C. Road, Calcutta 700 009, India

(*Received* 15 *August* 1993; *Accepted* 19 *August* 1993)

An analytical model has been developed for investigating the temporal variation of the marine boundary layer over the Indian Ocean Sector of the Southern Ocean surrounding Antarctica. The model inputs were provided from sparse data collected during the Indian Expeditions to Antarctica. In spite of the basic simplicity of the model, and the inadequacy of the available data, the model reproduces the boundary layer characteristics fairly well. These results have been presented in this paper.

Key Words: Boundary Layer; Southern Ocean; Antarctica Expedition

Introduction

The Southern Ocean, encircling the ice covered continent of Antarctica and comprising of the southernmost sectors of the Atlantic, Indian and Pacific Oceans, is a data sparse region of the world. Although in recent years satellite coverages have provided much meteorological and oceanographic data, the lack of a suitably dense network of *in situ* observations from these regions is being felt keenly, because of the overwhelming need for a "ground truth" database for calibrating and interpreting satellite data.

The role of the Southern Ocean and Antarctica in maintaining the climate and weather of not only the Southern Hemisphere, but the whole world has long been realised. In particular, the complex interactive feed-back processes between the atmosphere, the Antarctic ice cap, the surrounding seasonally varying sea-ice extent and the ocean waters need to be well understood for a number of reasons: operational weather forecasting over Antarctica and the Southern Hemisphere, providing accurate boundary conditions for theoretical models to simulate future climatic trends are some of the problems that need attention.

In this connection, a study of the atmospheric boundary layer, within which the atmosperic-cryosphere-ocean interactions are predominantly observed, forms an extremely interesting area of investigations.

In this paper, we have made an attempt to study the temporal evolution of the marine boundary layer (MBL) over the Indian Ocean sector of the Southern Ocean, confining our study to the region between 60°S-70°S at locations dictated by the ship route followed during the annual Indian Scientific Expeditions to Antarctica.

As far as the present authors are aware, there has been hardly any attempt at studies of the boundary layer in this particular sector of the Southern Ocean. However, studies of the atmospheric boundary layer over the Antarctic continent has been taken up by several authors earlier[1-8]. Similar studies over the Arctic and the colder regions of the Northern Hemisphere are more numerous.

It must be realised that in any theoretical investigation, the inadequacy of available data both for realistic inputs to the model and for subsequent validation of the model output, is a very serious limitation. For the present study, it was necessary to fall back on the data collected and published by the India Meteorological Department for the period January 1985 to February, 1989 (*Fourth Wintering Expedition to Eighth Summer Expedition*) (IMD, (1992), *Indian Expeditions to Antarctica,* Vols III and IV). The dataset consists of a limited number of radiosonde observations on the expedition ship and Antarctica, in addition to surface observations.

The layout of the paper begins with a description in the next section of an analytical model. The model was run with inputs from ship data combined with surface marine observations.

Model results for some case studies have been discussed briefly and illustrated graphically in the third section. In the concluding section, we have discussed some limitations of the available data, which must be removed as early as possible if such pursuits are to be continued seriously in future.

An Analytical Model for Boundary Layer Evolution

Firstly, we assume that the MBL is characterised by a potential temperature, $\theta_s = \theta_s(x, t)$ and mixing ratio $q_s = q_s(x, t)$ with temperature and moisture jumps $\Delta_\theta(x, t)$ and $\Delta_q(x, t)$ at the MBL top $h(x, t)$. Overlying the boundary layer we assume a layer of stable, free atmospheric air with constant lapse rates γ_1 (for temperature) and γ_2 (for moisture). This boundary layer structure has been used previously by several authors[2,9-13].

Neglecting advection, subsidence and any external forcings, the equations governing the evolution of θ_s and q_s are respectively:

$$\frac{\partial \theta_s}{\partial t} = -\frac{\partial}{\partial z}(\overline{w'\theta'}); \qquad \ldots (1)$$

and

$$\frac{\partial q_s}{\partial t} = -\frac{\partial}{\partial z}(\overline{w'q'}), \qquad \ldots (2)$$

which on integration with respect to z between $z = 0$, $h(x, t)$ yields

$$h\frac{\partial \theta_s}{\partial t} = \overline{w'\theta'_s} - \overline{w'\theta'_h} \qquad \ldots (3)$$

$$h\frac{\partial q_s}{\partial t} = \overline{w'q'_s} - \overline{w'q'_h} \qquad \ldots (4)$$

In (3) and (4), $\overline{w'\theta}$, $\overline{w'q'}$ are the vertical components of the heat moisture flux and subscripts s and h denote their values at surface level and at $z = h(x, t)$ respectively.

The potential temperature and humidity jumps, Δ_θ and Δ_q will evolve with time, increasing or decreasing because of

(a) Entrainment of the air from aloft the MBL, and

(b) Warming or cooling of the MBL due to changes in the surface (ocean) conditions

Mathematically, these can be formulated as the rate equations[10] for Δ_θ and Δ_q

$$\frac{d}{dt}\Delta_\theta = \gamma_1 \frac{dh}{dt} - \frac{\partial\theta_s}{\partial t} \qquad \ldots (5)$$

and

$$\frac{d}{dt}\Delta_q = -\gamma_2 \frac{dh}{dt} + \frac{\partial q_s}{\partial t}. \qquad \ldots (6)$$

The MBL top will evolve due to these processes in accordance with[9,11].

$$\Delta_\theta \frac{dh}{dt} = -\overline{w'\theta'_h} \qquad \ldots (7)$$

and

$$\Delta_q \frac{dh}{dt} = -\overline{w'q'_h} \qquad \ldots (8)$$

Using (7), (3) in (5) and (8), (4) in (6), we have

$$\left(\frac{d}{dh}\Delta_\theta - \gamma_1 + \frac{\Delta_\theta}{h}\right)\frac{dh}{dt} = \overline{w'\theta'_s} \qquad \ldots (9)$$

and

$$\left(\frac{d}{dh}\Delta_q + \gamma_2 - \frac{\Delta_q}{h}\right)\frac{dh}{dt} = -\frac{\overline{w'q'_s}}{h}. \qquad \ldots (10)$$

Combining (7) and (8) we find

$$-\overline{w'q'_h} = \frac{\Delta_q}{\Delta_\theta}(-\overline{w'\theta'_h}). \qquad \ldots (11)$$

Following Tennekes and Driedonks[14], we close the system of equations (3), (4), (7), (8) by

$$-w'\,\theta'_h = A\,w'\,\theta'_s + \frac{Bu_*^3\,\theta_s}{gh}, \qquad \ldots (12)$$

where $A = 0.2$, $B = 2.5$ are constants, u_* is the friction velocity and g is the acceleration due to gravity. The first term in the right hand side of (12) incorporates the effect of mechanical turbulence while the second, considers the effect of convection. With $B = 0$, $A = 1$, Ball's closure relation is recovered. The accepted value of B is still debated and ranges between 2.5 and 5[15].

The Monin-Obukhov length L is

$$L = -\frac{u_*^3\,\theta_s}{k_0\,g\,(\overline{w'\,\theta'_s})}, \qquad \ldots (13)$$

where $k_0 = 0.4$ is the von Karman constant.

Thus, with (12) and (13), $\overline{w'\,\theta'_h}$ can be expressed by

$$-\overline{w'\,\theta'_h} = \frac{w'\,\theta'_s}{h}\,[Ah - Bk_0\,L], \qquad \ldots (14)$$

whence (9) and (10) become the following first order ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}h}\Delta_\theta + \Delta_\theta\left(\frac{1}{h} + \frac{1}{Ah - Bk_0\,L}\right) = \gamma_1 \qquad \ldots (15)$$

$$\frac{\mathrm{d}}{\mathrm{d}h}\Delta_q - \frac{\Delta_q}{h} = -\dot{\gamma}_2 + \Delta_\theta \cdot \frac{\overline{w'\,q'_s}}{\overline{w'\,\theta'_s}} \cdot \frac{1}{Ah - Bk_0\,L}. \qquad \ldots (16)$$

In deriving (16) we have also used (11).

We mention that eq. (15) is exactly similar to the one obtained by Gryning and Batchvarova[15] (hereafter abbreviated GB) in their investigation of a coastal convective internal boundary layer. The solution of (15) is straight forward and is (GB):

$$\frac{\Delta_\theta}{\gamma_1 h} = \left[\frac{1}{1+2A}(Ah - Bk_0\,L)^{\frac{1+2A}{A}} + \frac{Bk_0\,L}{1+A}(Ah - Bk_0\,L)^{\frac{1+A}{A}} + \frac{A}{(1+A)(1+2A)}(-Bk_0\,L)^{\frac{1+2A}{A}}\right] \Big/ [Ah^2\,(Ah - Bk_0\,L)^{\frac{1}{A}}] \qquad \ldots (17)$$

It may be verified that:

$$(h/L) \to \pm 0, \quad (\Delta_\theta/\gamma_1\,h) = \frac{1}{2}$$

and

$$(h/L) \to \pm\infty, \quad (\Delta_\theta / \gamma_1 h) = A/(1+2A) \qquad \ldots (18)$$

Thus, a simpler form of (17) (see also, GB) is:

$$\frac{\Delta_\theta}{\gamma_1 h} = \frac{Ah - Bk_0 L}{(1+2A)h - 2Bk_0 L}. \qquad \ldots (19)$$

Using (19) in (16) we obtain the solution

$$\frac{\Delta_q}{\gamma_2 h} = \frac{\Delta q_0}{\gamma_2 h_0} + \ln\left(\frac{h_0}{h}\right) + C(t)\ln\left[\frac{(1+2A)h - 2Bk_0 L}{(1+2A)h_0 - 2Bk_0 L}\right], \qquad (20)$$

where h_0 is the boundary layer height at $t = t_0$ when $\Delta_q = \Delta_{q_0}$ and

$$C(t) = \frac{\gamma_1}{(1+2A)\gamma_2} \cdot \frac{\overline{w'q'_s}}{\overline{w'\theta'_s}}. \qquad \ldots (21)$$

Next we obtain an equation for $h(x, t)$ by combining (7) and (14) and (19):

$$\frac{h^2}{(1+2A)h - 2Bk_0 L}\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{1}{\gamma_1}\overline{w'\theta'_s}$$

Integrating between $t = t_0$, $t_0 + \Delta t$ at which $h = h_0$ and $h(x, t)$ respectively, after assuming that $\overline{w'\theta'_s}$, γ_1 and L do not change during Δt, we get

$$h^2 + \frac{4Bk_0 L}{1+2A}h + 8\left(\frac{Bk_0 L}{1+2A}\right)^2 \ln\left[\frac{(1+2A)h - 2Bk_0 L}{(1+2A)h_0 - 2Bk_0 L}\right] + f(h_0) = 0 \qquad \ldots (22)$$

where,

$$-f(h_0) = h_0^2 + \frac{4Bk_0 L}{1+2A}h_0 + \frac{2(1+2A)}{\gamma_1}\overline{w'\theta'_s} \cdot \Delta t \qquad \ldots (23)$$

To compute $\overline{w'\theta'_s}$, $\overline{w'q'_s}$ in our model we use the conventional bulk expressions

$$\overline{w'\theta'_s} = C_h|\mathbf{V}|(\theta_{SST} - \theta_s),$$

$$\overline{w'q'_s} = C_e|\mathbf{V}|(q_w - q(\theta_s)), \qquad \ldots (24)$$

and

$$q_w = q_*(\theta_{SST}),$$

where $C_h = 1.3 \times 10^{-3}$ and $C_e = 1.5 \times 10^{-3}$

which are the transfer coefficients for heat and moisture and the fluxes in (24) are all expressed in kinematic units, q and q_* denote the mixing ratio and its saturation value at temperature θ.

Other surface parameters, such as, u_* and the roughness length z_0 are computed by

$$u_*=(\tau/\rho)^{1/2}, \quad \tau/\rho=C_D|\mathbf{V}|^2,$$

$$C_D=0.7\times 10^{-3} \text{ (for oceans)}$$

and

$$z_0=0.016\, u_*^2/g \text{ (cf Charnock, 1955)}^{16}. \qquad \ldots (25)$$

In (24) and (25), **V** denotes the 10m wind, ρ the mean density of air and τ the surface stress.

Our system of model equations are therefore, (24), (25), (22), (19), (20), (3) and (4).

It now remains to compute the wind velocity V. We use the well-known similarity relationship

$$u=\frac{u_*}{k_0}\left[\ln\left(\frac{z}{z_0}\right)-\psi_m\right] \qquad \ldots (26)$$

where

$$\psi_m=\begin{cases}\ln\left[\left(\frac{1+x}{2}\right)^2\left(\frac{1+x^2}{2}\right)\right]-2\tan^{-1}x+\frac{\pi}{2}, & \text{for } L<0 \\ -5z/L, & \text{for } L>0\end{cases} \qquad \ldots (27)$$

(cf. Businger *et al.*, 1971[17])

z is taken as 10 m for the present study and in (27),

$x=(1-16z/L)^{1/4}$

The computation proceeds as follows:

For the first time step, θ_s, V, θ_{SST} from *r/s* and surface observations are provided as inputs to the model.

1. All surface quantities $\overline{w'\theta'_s}$, $\overline{w'q'_s}$, u_*, z_0, L are computed using (24), (25), (13)

2. The MBL height h is computed from (22) with h_0 prescribed from the *r/s* profiles of θ in the first time step.

3. Δ_θ and Δ_q are computed using (19) and (20) with Δ_{q_0} prescribed at $t=t_0$ from the *r/s* data.

4. $\overline{w'\,\theta'_h}$, $\overline{w'\,q'_h}$ are computed from (14) and (11) respectively.

5. $\theta_s\,(t+\Delta t)$ and $q_s\,(t+\Delta t)$ are computed from the finite differenced forms of (3) and (4):

$$\theta_s(n+1)=\theta_s(n)+\Delta t\;\frac{w'\,\theta'_s-w'\,\theta'_h}{h}$$

and

$$q_s\,(n+1)=q_s(n)+\Delta t\;\;w'\,q'_s+\Delta_q\cdot w'\,\theta'_h/\Delta_\theta$$

For the present study, we have used $\Delta t=600$s.

6. $u(n+1)$ is computed using (26) and (27).

7. The newly computed values of u, θ_s, q_s, Δ_θ, Δ_q, and h are used in the computations repeated from Step 1.

Discussion of the Model Results

The model developed in the foregoing sections has been used to make four case studies for 17, 18 December, 1986, 22 December, 1988, and 2 March, 1989. The rationale behind the choice of these dates has been the availability of radiosonde and corresponding surface observations providing the Sea Surface Temperature (SST). However, we also find that this choice of data allows us to make a comparative study of the latitudinal and seasonal variations of the MBL parameters. The position of the expedition ship on these four days has been given in Table I. It should be remembered that December in the Southern Hemisphere, is a summer month while March is usually classified as autumn, particularly in and around Antarctica.

All the radiosonde observations were taken at 12 GMT. With these observed values of the SST, air temperature, dew point temperature wind speed and direction and mixing ratio, the model was integrated for 3 hours with 10 minute time steps, as mentioned earlier.

The temporal variation of the surface heat flux, $\overline{w'\,\theta'_s}$ has been shown in Fig. 1. We find that on 22 Dec '88 and 2 Mar '89, the surface heat flux is posi-

Table I

Position of the expedition ship on the four case study days

Date	Position
17 December 1986	62.2 S/23.2 E
18 December 1986	65.1 S/14.8 E
22 December 1988	67.4 S/18.0 E
2 March 1989	70.1 S/12.0 E

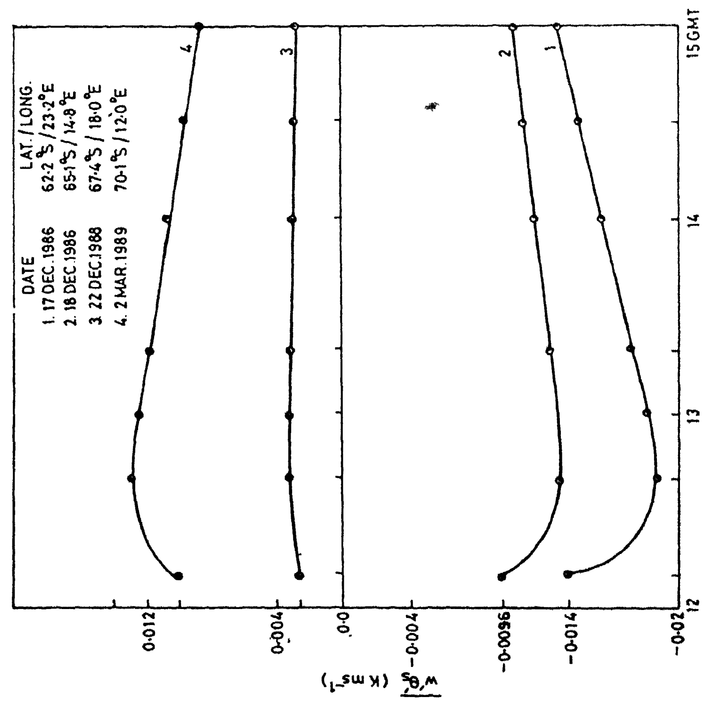


Fig 1 Variation of surface heat flux $\overline{w'\theta'_s}$ with time

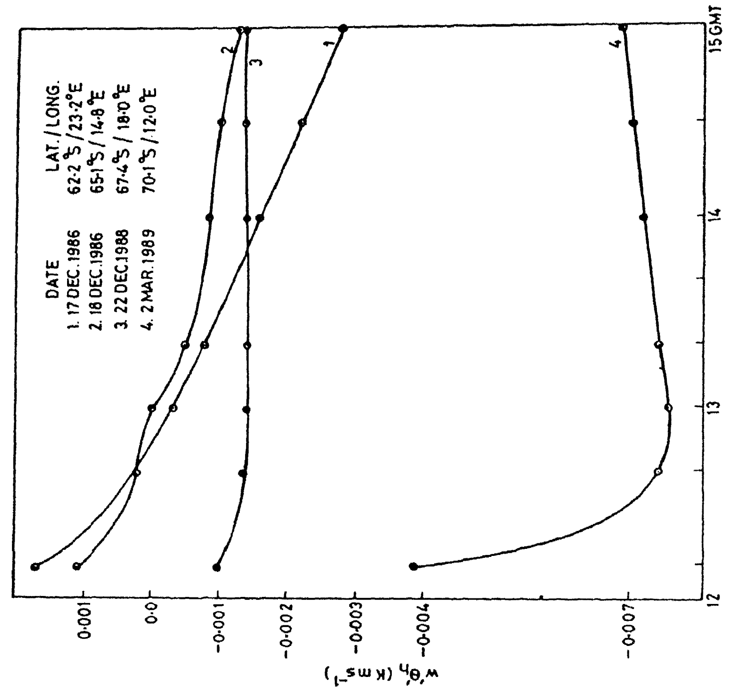


Fig 2 Variation of heat flux at boundary layer top with time

tive, indicating a loss of heat from the ocean surface. Taking into account the proximity of the locations on these days to the Antarctic ice shelf, this loss of heat can obviously be attributed to the low air temperatures over a relatively warmer ocean surface in these regions falling within the 67°S-71°S belt. On the other hand, at the more northern latitudinal belt of 62°S-65°S (case studies for 17, 18 Dec 1986), the surface heat flux is negative, indicating a downward transfer of heat at the ocean surface.

Fig 2 illustrates the temporal variation of the heat flux $\overline{w'\theta'_h}$ at the MBL top, for all four days. Again one finds the emergence of two distinct latitudinal regimes: one between 62°S-65°S and another between 67°S-71°S. Thus, on 22 Dec, '88 and 2 Mar '89, $\overline{w'\theta'_h}$ continues to remain negative all throughout the three hours of model integration indicating a downward transport from the MBL top. Fig. 2 also shows that on the Antarctic coast (71.1°S, 12°E), there is a very rapid decrease of $\overline{w'\theta'_h}$ for the first forty minutes of model integration indicating a rapid cooling of the MBL top. After this, the downward transport of $\overline{w'\theta'_h}$ proceeds at a much slower rate. In contrast, on both 17 Dec '86 and 18 Dec. '86, corresponding to the relatively warmer latitudinal belt of 62°S-65°S, the temporal variation of $\overline{w'\theta'_h}$ is from a positive (upward transport) at 12 GMT to a negative value (downward transport) at 15 GMT. The heat flux reaches zero a little prior to 13 GMT on 17 Dec. while on 18 Dec., zero heat flux at the MBL top occurs a little after 13 GMT.

The surface moisture flux $\overline{w'q'_s}$ comparatively speaking, exhibits simpler variation patterns, both latitudinally and temporally, as can be seen from Fig. 3. The figure also shows that $\overline{w'q'_s}$ increases as one moves Southwards towards Antarctica, possibly because of the increasing air-sea temperature difference.

The moisture flux $\overline{w'q'_h}$ (Fig. 4) at the MBL top varies prominently in the relatively warmer 62°S-65°S belt. Again, as one moves towards Antarctica, $\overline{w'q'_h}$ becomes very small and remains nearly constant with time. This seems to be compatible with the fact that the air temperature in and around Antarctica remains so low that hardly any moisture is retained and consequently, moisture transport is greatly reduced.

In Fig. 5, the variation of the boundary layer height, h, as computed from the model has been shown. At the warmer latitudes, 62°S-65°S, the MBL height tends to decrease with time, reaching approximately steady values of 285m (17 Dec '86) and 440m (18 Dec '86) at 15 GMT. In the colder latitudes, 67°S-71°S, the MBL height increases with time at a slow rate, rising to steady values of 631m (22 Dec. '88) and 691m (2 Mar. '89) at 15 GMT.

Finally, in Fig. 6, the observed vertical profiles of the potential temperature have been shown for all the four days. The dotted lines indicate the height above sea level of the first temperature inversion as revealed by the soundings. Assuming this level is close to the actual MBL height, we find that the computed MBL heights agree fairly well with these.

The main conclusions are:

i) An analysis of this kind is handicapped by lack of information on upper air winds.

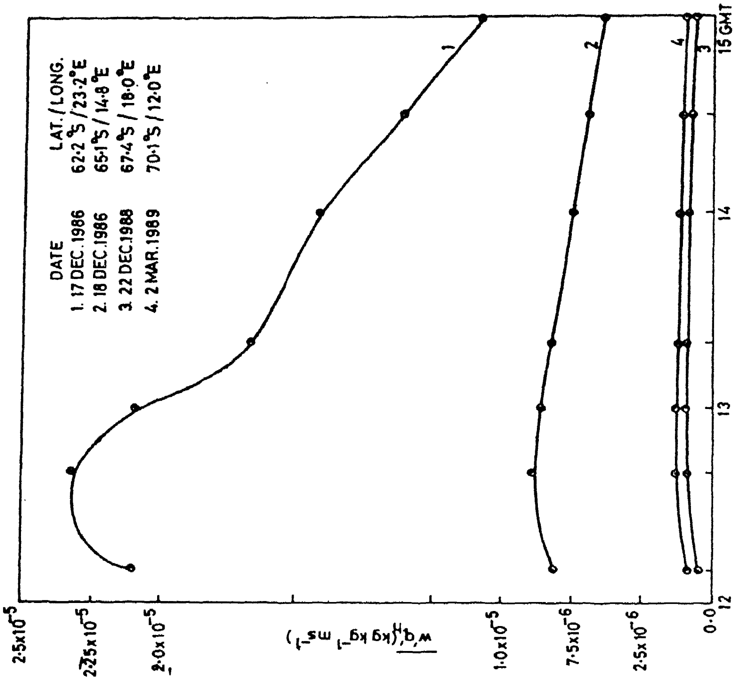


Fig 4 Variation of moisture flux at boundary layer top with time

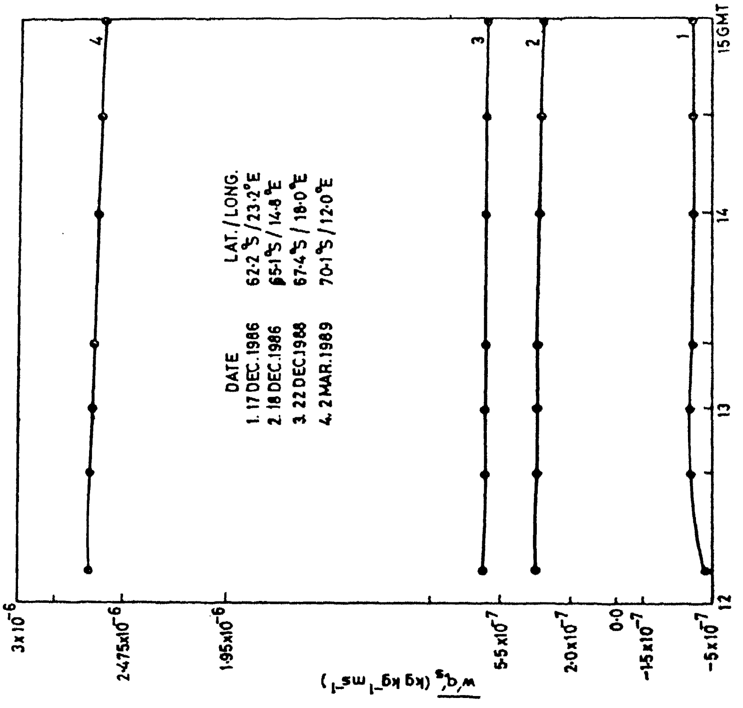


Fig 3 Variation of surface moisture flux $\overline{w'q'_s}$ with time

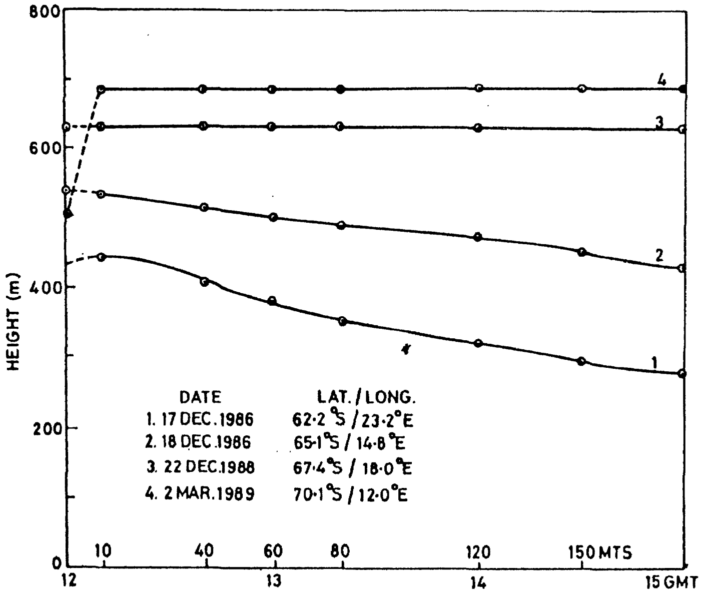


Fig 5 Variation of boundary layer height

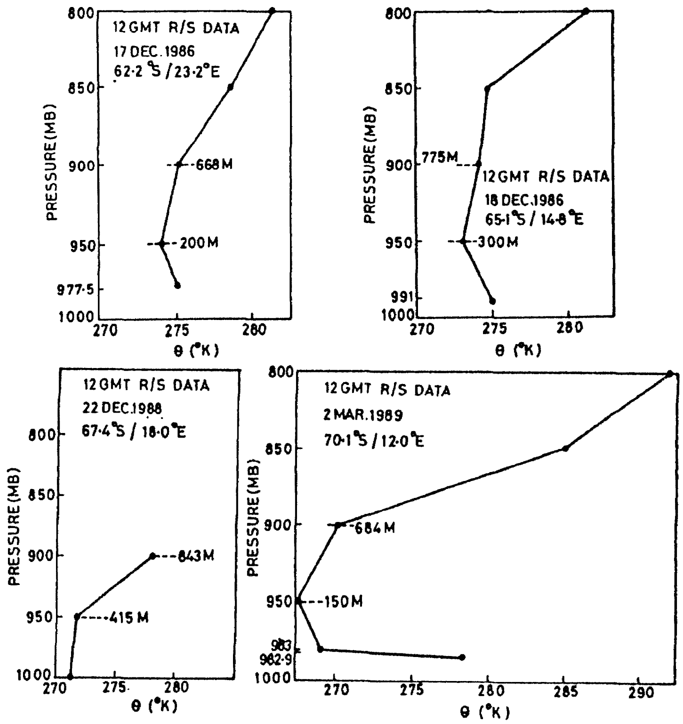


Fig 6 12 GMT potential temperature profiles

ii) On 22nd December 1988 and 2nd March 1989 the surface heat flux is positive, indicating a loss of heat from the ocean surface.

iii) Such a loss of heat is attributable to the low air temperatures over a relatively warmer ocean surface in these region.

vi) Heat flux continues to be negative indicating a rapid cooling of the MBL top.

v) The surface moisture flux also exhibits variation patterns.

Acknowledgement

The author gratefully acknowledges the Department of Science & Technology, Govt. of India for providing research grants and the India Meteorological Department for providing the relevent datasets.

Dr G K Sen and Professor D K Sinha, Centre for Atmospheric Sciences, Calcutta University are being gratefully acknowledged for many fruitful discussions and help in connection with this study.

Finally, the Department of Ocean Development, Govt. of India is also acknowledged for all their help during Dr Chatterjee's participation in the tenth Indian Scientific Expedition to Antarctica, which was instrumental in motivating the present investigation.

References

1 G H Liljequist *Energy Exchange Over an Antarctic Snow Field: Norwegian-British Swedish Expdn.* 1949-1952, *Scient. Results. II* Norsk Polarinstitut Oslo (1957)
2 F K Ball *Antarctic Met* (1960) 9-16 Pergamon New York
3 F K Ball *Q J R met Soc* **86** (1960) 483-494
4 P C Dalrymple, H Lettau and S Wollaston *Studies in Antarctic Meteorology: Antarctic Res Sr* **9** (Ed M J Rubin) Washington D.C. (1966) 13-57
5 W Schwerdtfegger *Weather and Climate of the Antarctic* Elsevier New York (1984) 261 pp
6 T Parish *Mon Wea Rev* **110** (1982) 84-90
7 J C King *Antarctic Sci* **1** (1989) 169-179
8 J C King *Q J R met Soc* **116** (1990) 363-378
9 D K Lilly *J atoms Sci* **25** (1968) 1092-1099
10 H Tennekes *J atoms* **30** (1973) 558-581
11 H Okland *ECMWF Tech Rep* No 16 (1979) 29 pp
12 E B Kraus and L D Leslie *J atoms Sci* **39** (1982) 2760-2772
13 W Brutsaert *Bound Layer Met* **39** (1987) 283-300
14 H Tennekes and A G M Driedonks *Bound Layer Met* **20** (1981) 515-531
15 S E Gryning and E Batchvarova *Q J R met Soc* **116** (1990) 187-203
16 H Charnock *Q J R met Soc* **81** (1955) 639.
17 J A Businger, J C Wyngaard, Y Izumi and E F Bradley *J atmos Sci* **28** (1971) 181-189

TERRAIN VARIANCE SPECTRA FOR INDIAN WESTERN GHATS

K SRINIVASAN AND N RAMANATHAN
Centre for Atmospheric Sciences, Indian Institute of Technology
Hauz Khas, New Delhi 110 016, *India*

(*Received* 3 *September* 1993; *Accepted* 1 *October* 1993)

The terrain height variance spectra for the Pune region, including the Western Ghats, are calculated with the available topographic map along three different cross sections. The required minimum grid size for mesoscale models to adequately represent terrain forcing, without resorting to sub-grid scale parameterization, is determined in this study Calculated values show that over 85% of the terrain variations are accounted for, if the chosen grid interval is of 4km. The terrain variances are found to vary proportional to the square of the wave length from the calculated spectra.

Key Words: Spectral Studies; Grid Resolution

Introduction

The distribution of terrain height variance with wavelength is the variance spectrum. This spectrum gives us an idea about the dominant wavelengths of terrain variation, which will be useful for finding the optimum horizontal gride size for numerical models of orographic flow. The purpose of this present study is to determine the optimum grid size for the area taking into account terrain variances. For the purpose of the study, the eastern slope of the Western Ghats comprising Poona and its environs is chosen. This area is prone to thunderstorms and mesoscale weather developments.

The results of previous investigators vary considerably depending on geomorphology of the studied region. For example, Bretherton[1] found the terrain variance, for northern Wales, to be proportional to $\lambda^{3/2}$ for $\lambda < 30$km. For West Central Virginia in the U.S.A., Pielke and Kennedy[2] have shown terrain height variance to increase as λ^2 for short wave lengths. Young and Pielke[3] found the dependence as λ^1 for three different cross sections in the Colorado area. Since terrain variations are dissimilar throughout the world, it is necessary to study each region for application in NWP models.

Based on the above studies, this study determines terrain variance spectra for the above mentioned region, using a Fast Fourier Transform (FFT) method. The methodology and data are described in the next section. In the third section, the results are discussed. Summary and conclusion form the last section.

Data

Detailed fine resolution topographic maps are usually not available. The available digitized topographic data are of coarse resolution ~ 15km, and not suitable for this study. In this study, a published topographic map with a contour interval of 250' is used (Atmanathan[4]). This is shown in Fig. 1. The terrain

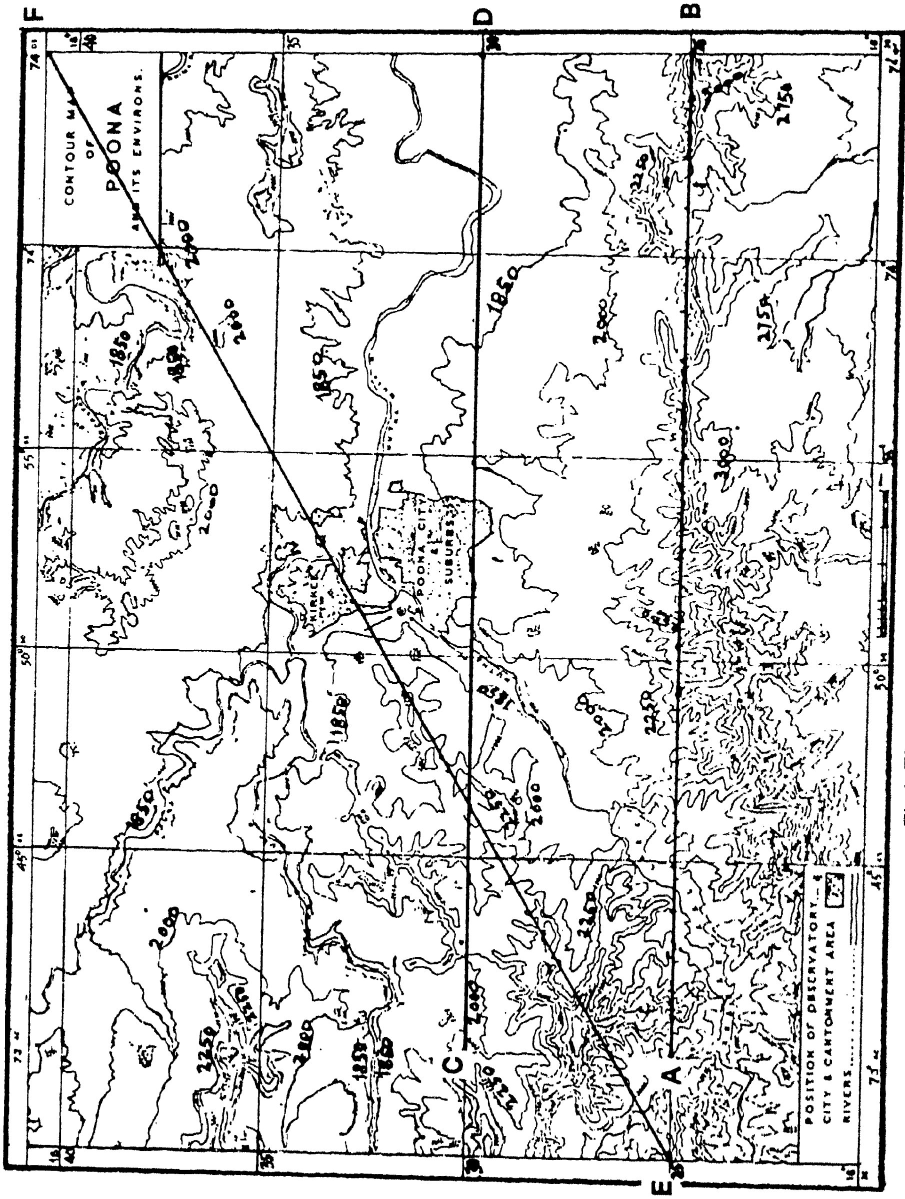


Fig 1 Elevation contour map of the study region.

elevations were evaluated manually from Fig. 1 along three different cross sections as shown by the lines AB, CD and EF. For each cross section 80 equally spaced intervals were taken along these lines tq evaluate the terrain heights. A one-dimensional variance spectra for the data set were computed using the FFT algorithm. For computational purposes, the deviations of terrain elevations from the average value along each cross sections are obtained. The resulting height series was transformed to wave number space (k) and the terrain variances were calculated. The calculated distribution of variance as a function of wavelength (λ) or wave number (k) are presented in Figs 2 to 5. Following

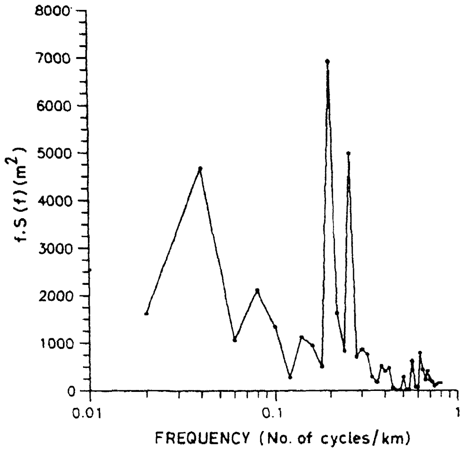


Fig 2 Spectral density (S) multiplied with its frequency (f-cycles/km) is plotted as a function of frequency and wavelength (λ), for the cross section AB

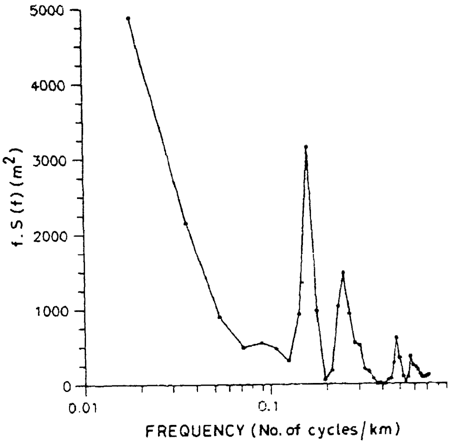


Fig 3 Same as fig 2 but for EF.

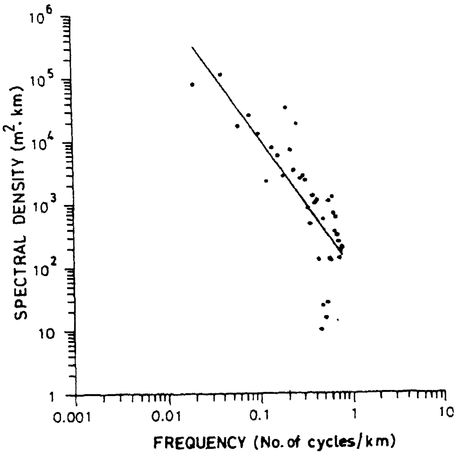


Fig 4 Terrain height variance spectra for the cross section AB. Spectral density (S) is plotted as a function of frequency and wavelength in log-log scale

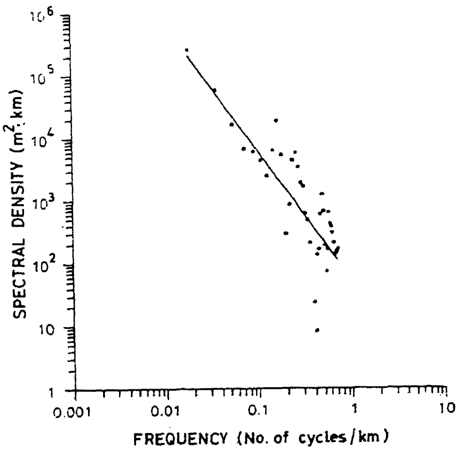


Fig 5 Same as Fig 4 but for EF.

Young and Pielke[3], a linear trend was also calculated for the terrain height series by the method of least squares and then subtracted from the terrain height series for each cross section. The calculated spectra for this series yielded similar results as those obtained with the series by subtracting the average value of the terrain series. For brevity, the calculated spectra using the first method are shown here.

Results and Discussion

The major use of terrain spectra is for determining the required spatial resolution in mesoscale numerical models. In such simulations the smallest atmospheric circulation that can be represented has a horizontal wave length equal to twice the minimum grid resolution in the model. Over rough terrain, this characteristic of the model also states that the smallest topographic feature which can be resolved is equal to twice the minimum grid interval. Thus, for example, if the grid spacing in the model is 10km, but 90% of the terrain irregularities are 5km or smaller, there is little hope of correctly simulating the atmospheric response to this topography. But, on the other hand, if the grid spacing was at 2km intervals and 90% of the variations are larger than 4km, the model should be capable of realistically representing the situation.

The results are presented in two different ways. In Figs 2 and 3 the contributions of variance are plotted as a function of wavelength. This shows the horizontal spacing of most typical terrain features, as well as the amount of variance. In an alternate representation, the terrain height variances are plotted against wavelength (λ) or wavenumber (k) in a logarithmic scale. A power law relationship between the variables such as $S = a \cdot \lambda^b$ was fitted by the method of least squares. This analytic form is dictated by the nature of the distribution of points in the diagonal. The power law dependence for atmospheric circulations is also well known from earlier studies (Vander Hoven[5]). In Figs 2 and 3 by plotting $f \cdot S(f)$ against $\log(f)$, (f is frequency with units No. of cycles per km, and $S(f)$ is the spectral density with units m^2 km associated with the frequency f). The low frequency portions of the spectra are expanded along the abscissa. Also, the ordinate for the high frequency portions are enhanced because the spectral density is multiplied by frequency. Another excellent quality is that the area under any portion of the curve continues to the porportional to the variance. Further, it gives us an idea about the dominant wavelengths towards the contribution of variance in the calculated spectra. From Fig. 2, the dominant wavelengths are 25km, 5km, and 4km over the cross section AB. Similar results were obtained for the cross section CD. Over cross section EF, from Fig. 3, the dominant wavelengths are 56km, 6km and 4km.

In Figs 4 and 5 the spectral density is plotted as a function of wavenumber on a logarithmic scale. The power law relationship between the variables (i.e.) spectral density (S) and wavelength (λ) in the form of $S = a \cdot \lambda^b$ was fitted by the method of least squares. The results are tabulated in Table I.

Table I

Parameters of the least squares best fit relation of the form $S = a \cdot \lambda^b$ for three terrain height variance spectra considered over Poona

[The tabulated parameters are coefficient a, the exponent b, the correlation coefficient (γ) between log(f) and log(S) and standard.deviation S_b of the exponent.]

Data set	a	b	γ	S_b
Cross section AB	96.28	2.08	0.80	0.25
Cross section CD	12.00	2.02	0.85	0.20
Cross section EF	53.72	2.07	0.75	0.29

The value of the exponent b was used to discuss the choice of horizontal grid spacing for mesoscale models. From Table I, the value of the exponent b is almost same in all the three cross sections considered, where as the coefficient a varies between cross sections because of differing geographic coverage.

Confidence intervals may be computed for the true value of the exponent b using $F = \frac{(B-b)^2}{S_b^2}$, where F has the degrees of freedom 1 and n-2, n being the number of harmonics in the spectra (Kreyszig[6]). A value of $\alpha = 0.05$ was selected for the computation of confidence intervals. The confidence intervals for the true value of the exponent were $B = 2.08 \pm 0.5$ for the cross section AB, $B = 2.02 \pm 0.4$ for the cross section CD and $B = 2.07 \pm 0.58$ for the cross section EF.

Similar studies on other regions of the world indicate that the spectral densities varies proportional to λ^1 or λ^2 with corresponding variations in the coefficient a (Pielke and Kennedy[2] and Young and Pielke[3]). However, our study has shown that the terrain variance varies as λ^2 as shown in Table I. Since the terrain heights and the configuration are dissimilar throughout the world, the obtained values are reasonable and provides a quantitative measure on terrain smoothness.

For nonlinear mesoscale models, using a terrain following co-ordinate system, Pielke[7] has shown the term containing subgrid scale terrain fluctuations can be neglected in grid volume averaged conservation equations if the subgrid scale terrain variance is small compared to grid resolvable terrain variance. We can find a maximum grid size for the model, so that the subgrid scale terrain variance can be neglected. For terrain height variance spectra of the form $S = a \cdot \lambda^2$, the ratio of the subgrid terrain height variance to model resolved terrain height variance is

$$\int_{1/2\delta x}^{1/2\Delta x} ak^{-2}\mathrm{d}k \qquad \int_{1/2\Delta x}^{1/n\Delta x} ak^{-2}\mathrm{d}k,$$

where $k = 1/\lambda$, $2\delta x$ is the shortest wavelength in the measured spectra, $2\Delta x$ is the shortest wavelength that can be resolved by the numerical model and $n\Delta x$ is the model domain length. This ratio turns out to be 0.15 for a 50km domain, (from the spectra) $2\delta x = 1.25$km and model grid size 4km. Thus, numerical models of mesoscale orographic flows over Poona region, requires a minimum horizontal grid spacing of 4km to resolve adequately the terrain without resorting subgrid scale terrain variance parameterization.

Conclusions

One dimensional terrain height spectra for three different cross sections over a region in and around Poona with western ghats were calculated. The calculated results show that a horizontal gride spacing of 4km is necessary to resolve the majority of terrain height variations without requiring a subgrid scale parameterization for the terrain height variances. A power law fit for the obtained spectra show that the terrain variances are proportional to λ^2. The presented results may serve as a guide for choice of grid interval in Numerical Weather Prediction models.

Acknowledgement

The computations were made on an ICL-3980 computer system installed at IIT-Delhi.

References

1 F D Bretherton *Q J R Met Soc* **95** (1969) 213-243
2 R A Pielke and E Kennedy *Rep-UVA-ENVSCI-MESO*-1980-1 (1980)
3 G S Young and R A Pielke *J atmos Sci* **40** (1983) 2555-2560
4 S Atmanathan *IMD Scient Note* **IV** (1931) 101-117
5 I Vander Hoven *J Met* **14** (1957) 160-164
6 E Kreyszig *Wiley Eastern Limited* (1992) p 988.
7 R A Pielke *Advances in Geophysics Academic Press* **23** (1981) 185-344

A NUMERICAL SIMULATION STUDY OF THE ROLE OF SEA SURFACE TEMPERATURE IN TROPICAL CYCLONE EVOLUTION

D V BHASKAR RAO AND S S V S RAMAKRISHNA

Department of Meteorology & Oceanography, Andhra University,
Visakhapatnam-530 003, *India*

(*Received* 15 *June* 1993; *Accepted* 5 *July* 1993)

The evolution of tropical cyclone circulation from an initial balanced vortex is studied in relation to the sea surface temperature (SST). A six-level axi-symmetric primitive equation model in cylindrical coordinates incorporating the Arakawa-Schubert cumulus parameterization scheme has been used. The SST values have been varied from 298 K to 301 K at one degree interval and the corresponding results indicate that tropical cyclone intensity increases with increase of SST and a minimum SST of 299 K is necessary for the model vortex to transform into a mature cyclone.

Key Words: Numerical Simulation; Sea Surface Temperature; Tropical Cyclone; Evolution; Arakawa-Schubert Cumulus Parameterization Scheme; CISK Mechanism

Introduction

Numerical modelling of tropical cyclones started with the early attempts of Kasahara[1] and Syono[2]. Many modelling studies followed with the development of the concept of CISK mechanism[3], defining the cooperative interaction of cumulus scale and large-scale circulations, which has been the basis of the parameterization of cumulus convection[4-6]. Axi-symmetric models have been designed to derive salient features of the tropical cyclone evolution and are largely successful[7-12]. Anthes[13] gave a good review of the tropical cyclones. Many of the numerical modelling experiments are confined to the simulation of the evolution of cyclones. The pioneering work of Palmen[14] has indicated that the tropical cyclones form over warm tropical oceans where the SST are greater than 26°C and are supported by Wendland[15]. Some attempts have been made to study the development of tropical cyclones in relation to sea surface temperatures using numerical models[16-20]. These model sensitivity studies supported that warm SST favour the intensification of cyclones with a threshold value of 26°C.

The cumulus parameterization scheme proposed by Arakawa and Schubert[6] (A-S scheme) is based on well-formulated physical processes of cloud formation and with a closure hypothesis dependent on the concept of quasi-equilibrium assumption between the cloud generating kinetic energy from large scale forces and the environmental stabilization due to clouds. Wada[11] first attempted the simulation of tropical cyclone using a simplified A-S scheme in which the mixed layer moist static energy (Hm) varies with an assumed constant heat supply. Bhaskar Rao[12] modified Wada's model by incorporating the computation of Hm. In this study, an attempt has been made to study the evolution of the cyclonic circulation in relation to SST values 298, 299, 300 and 301 K using an axi-symmetric primitive equation model in order to study the changes in the evolution at different SSTs and to assess the model's performance in validating the threshold value of SST at 299 K.

Description of the Model

The model used in this study is an axi-symmetric primitive equation model written in r, θ, z, coordinates. The main equations of the model are as follows:

Equation for the radial wind V_r

$$\frac{\partial V_r}{\partial t} = -V_r\frac{\partial V_r}{\partial r} - W\frac{\partial V_r}{\partial z} + \left[f + \frac{V_\theta}{r}\right] V_\theta - \theta\frac{\partial \Phi}{\partial r} + \kappa\left[\nabla_1^2 - \frac{1}{r^2}\right] V_r + \frac{1}{\bar{\rho}}\frac{\partial \tau_r}{\partial z} \qquad \ldots (1)$$

Equation for the tangential wind V_θ

$$\frac{\partial V_\theta}{\partial t} = -V_r\frac{\partial V_\theta}{\partial r} - W\frac{\partial V_\theta}{\partial z} + \left[f + \frac{V_\theta}{r}\right] V_r - \theta\frac{\partial \Phi}{\partial r} + \kappa\left[\nabla_1^2 - \frac{1}{r^2}\right] V_r + \frac{1}{\bar{\rho}}\frac{\partial \tau_\theta}{\partial z} \qquad \ldots (2)$$

Equation for Potential Temperature (Thermodynamic equation)

$$\frac{\partial \theta}{\partial t} = -V\mathrm{r}\frac{\partial \theta}{\partial r} - W\frac{\partial \theta}{\partial z} + \frac{1}{\bar{\rho}\bar{\Phi}}\left[M_c\frac{\partial s}{\partial z} + D(s_c - s - L\hat{l}_c)\right] + \frac{L}{\bar{\Phi}}C + \kappa_\theta\nabla_1^2\,\theta \qquad \ldots (3)$$

Equation for mixing ratio of water vapour

$$\frac{\partial q}{\partial t} = -V_r\frac{\partial q}{\partial r} - \omega\frac{\partial q}{\partial z} + \frac{1}{\bar{\rho}}\left[M_c\frac{\partial q}{\partial z} + D\,(q_c - q + \hat{l}_c)\right] - C + \kappa_q\nabla_1^2\,q \qquad \ldots (4)$$

Equation for continuity of mass

$$\frac{\partial}{\partial z}\bar{\rho}\,\omega = -\frac{1}{r}\frac{\partial}{\partial r}\bar{\rho}\,r\,V_r \qquad \ldots (5)$$

Hydrostatic equation

$$\frac{\partial \Phi}{\partial z} = -\frac{g}{\theta}, \qquad \ldots (6)$$

where

$$\Phi = C_p\left(\frac{P}{P_0}\right)^{R\;C_p}, \qquad \ldots (7)$$

$$\nabla_1^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}.$$

Explanation for the symbols used in given in Table I. Over bars denote functions of height alone and subscript c denotes values in the cloud. Over cap^ denotes values at the clould top level. The cumulus convection effects are included as 3rd and 4th terms in the eq. (3) and (4).

Table I

Description of the symbols

V_r	— radial wind component
V_θ	— tangential wind component
W	— vertical velocity
θ	— potential temperature
q	— mixing ratio of water vapour
l	— mixing ratio of the liquid water
p	— pressure
p_0	— 1000mb
ρ	— density of the air
f	— coriolis parameter
g	— acceleration due to gravity
R	— gas constant for air
L	— latent heat of condensation
C_p	— specific heat of air at constant pressure
S	— dry static energy
$M_B(Z)$	— total vertical cloud mass flux at level Z
D	— total detrainment of cloud mass flux at level Z
C	— large scale condensation per unit time and unit mass.
K	— horizontal eddy viscosity coefficient
K_θ	— horizontal eddy thermal diffusivity
K_q	— horizontal eddy moisture diffusivity
*mps	— meters per second.
*cmps	— centimeters per second.
*hr	— hour/hours.
*km	— kilometers.

$$\tau_r = \bar{\rho}\theta\frac{\partial V_r}{\partial z}, \quad \tau_\theta = \bar{\rho}\theta\frac{\partial V_\theta}{\partial z}. \qquad \ldots (8)$$

At the surface, $\tau_{rs} = \rho_s C_D \mid V_s \mid \theta_{rs}$, $\tau_{rs} = \rho_s C_D \mid v_s \mid \theta_{rs}$ ... (9)

C_D is the drag coefficient. Subscript s denotes surface values.

For computational accuracy ϕ and θ are divided into basic and perturbation components. The mixed layer equations are written as follows:

$$\frac{\partial S_M}{\partial t} = -V_r\frac{\partial S_M}{\partial r} + \kappa_H \nabla_1^2 S_M + F_s \qquad \ldots (10a)$$

and

$$\frac{\partial q_M}{\partial t} = -V_r \frac{\partial q_m}{\partial r} + \kappa_H \nabla_1^2 q_M + F_q, \qquad \ldots (10b)$$

where S_M and q_m are the mixed layer values of dry static energy and mixing ratio. F_S and F_q are the sensible and latent heat supplies from the ocean surface given by the bulk aerodynamic formulae

$$F_s = C_D \mid V_s \mid C_P(T_s - T_M)/Z_B \quad \text{and} \quad F_q = LC_D \mid V_s \mid (q_s - q_M)/Z_B, \qquad \ldots (11)$$

where

T_s = temperature at the sea surface
T_M = temperature of air at the sea surface in the mixed layer, and
q_s = mixing ratio at sea surface which is calculated as the saturated mixing ratio at the temperature of the sea surface and existing pressure at that point.

Suffix M denotes values in the mixed layer and z_B is the height of the mixed layer.

Boundary Conditions: To determine pressure field though mass continuity equation, vertical velocity W is taken as zero at $z = 0$ and at $z = z_{top}$ i.e., at the bottom and the top of the model atmosphere.

At the lateral boundaries

$$V_r = V_\theta = 0 \text{ at } r = 0 \text{ and } r = r_{\max} \qquad (12)$$

and

$$\frac{\partial \Phi}{\partial r} = \frac{\partial \theta}{\partial r} = \frac{\partial q}{\partial r} = 0 \quad \text{at } r = 0 \text{ and } r = r_{\max} \qquad \ldots (13)$$

$r_{\max}$ is the radius of the outer periphery of the computational domain which is taken as 2500km in this model.

Arakawa-Schubert Parameterization Scheme: In this model discretized version of Arakawa and Schubert scheme (AS scheme) is used. The cloud base is fixed at 0.5km and three types of clouds with their tops fixed at 4km, 8km and 12km are only possible because of the vertical resolution as shown in Fig. 1. For each cloud type the cloud base mass flux is defined at half integer levels and entrainment or detrainment takes place at the integer levels. The entrainment rate of each cloud type is calculated with the cloud tops fixed as mentioned earlier. At every time step the convection computations are performed. For the computation of entrainment rate and cloud work function the temperature, the moisture field and the moist static energy are used. A cloud can exist only if the entrainment rate is positive and the virtual static energy in the cloud exceeds the environment value at all levels. When these conditions are satisfied, the equation for the cloud base mass flux is solved. Details of this procedure are same as given in Lord *et al.*[21]

Finite difference Scheme: The vertical structure of the model is given in Fig. 1. The atmosphere is divided into five layers, each layer bound by half-integer. V_r, V_θ and ϕ are defined at integer levels while W, θ and q are defined at half integer levels. The layer between levels 1 and 1 ½ is the mixed layer. An

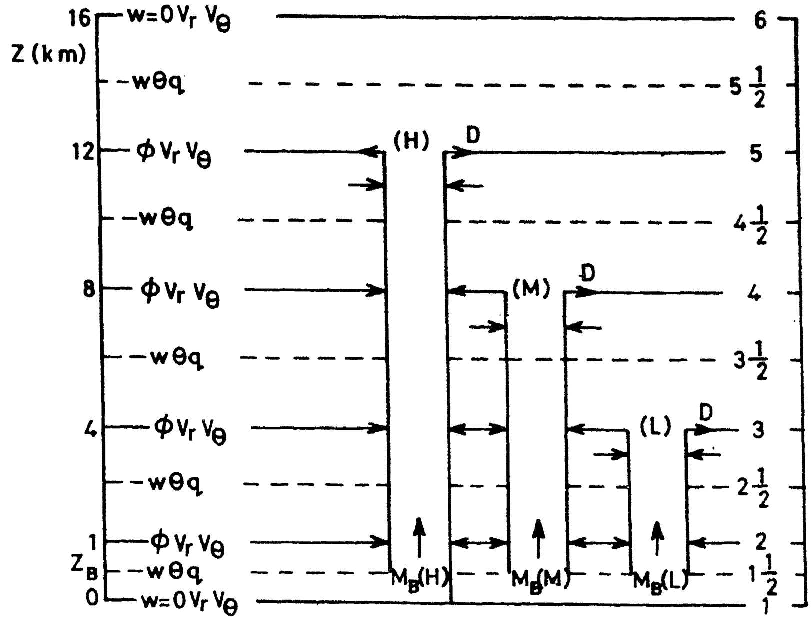


Fig 1 Vertical structure of the model

infinitesimal thin layer is assumed between $Z_{1\frac{1}{2}}$ level and the top of the mixed layer and S_M and q_M are taken as cloud base values. Equations of motion are applied at integer levels and the thermodynamic equation and the water vapour mixing ratio equation are applied at half integer levels. The horizontal computational domain is divided into 21 concentric rings (Table II). The six dependent variables are placed in staggered manner with V_r and V_θ defined at

Table II

Description of the grid points

Grid No.	Distance (km.)	Grid No.	Distance (km.)
1	20	12	443
2	40	13	538
3	60	14	651
4	80	15	787
5	102	16	951
6	128	17	1147
7	160	18	1382
8	198	19	1665
9	244	20	2004
10	298	21	2500
11	364		

integer points while W, θ, q and P are placed at half integer points. Upstream finite differencing is used for space differentiation. Time staggering is used for time derivative with θ, q and ϕ defined at $T\Delta T$ time intervals, while V_r, V_θ and W are defined at intermediate points as $(T+\frac{1}{2})\Delta T$. A time step of 300secs is used to avoid computational instability.

Initial Conditions

For all the experiments in the present study, the initial state is same and designed to have a weak balanced vortex in the wind field. This is obtained through a perturbation in the potential temperature field given by equation

$$\theta'(r_1, Z) = 0.16\left\{\cos\left(\frac{\pi}{r_0}\right) r + 1.0\right\} \sin \tag{14}$$

for $r < r_0$, where r = radius; r_0 = 200km; Z = height; Z_t = upper boundary of the model atmosphere.

The initial pressure is derived from the hydrostatic equation and the wind from the gradient wind equation. The above perturbation yields a weak balanced circulation with a maximum tangential wind of 8mps at 128km from the centre at the 0km and 1km levels. The tangential wind decreases linearly with height and becomes zero at the 16km level, the model top. The tangential wind at the surface is about 5mps between 60km and 160km and is zero beyond 198km. The radial winds are zero everywhere intially. The initial surface pressure is 1010mb beyond 221km and gradually decreases to 1008.5mb at the 10km radius. The potential temperature perturbation is maximum with an in-

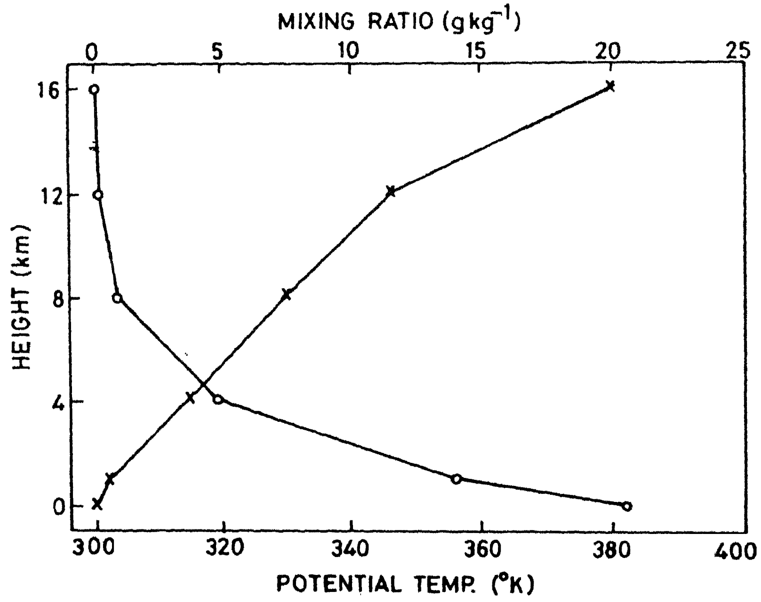


Fig 2 Vertical distributions of potential temperature (cross) and water vapour Mixing ratio (circle)

crease of 0.42K at 10km level. The initial vertical distributions of potential temperature and water vapour mixing ratio are the same as of the mean tropical atmosphere sounding of Jordan[22] and are shown in Fig. 2. The corresponding moist static energy, and saturation moist static energy are given in Fig. 3. This atmospheric state is conditionally unstable with lapse rates exceeding the moist adiabatic lapse rate upto 4km level and then moist adiabatic upto 8km and above 8km stability increases gradually.

The initial mixing ratio values are uniform horizontally. Values of dry static energy (S_m) and the mixing ratio (Q_m) in the mixed layer are computed from the surface values which gives the moist static energy in the mixed layer (H_m), an initial value of 84.5 cal/gm.

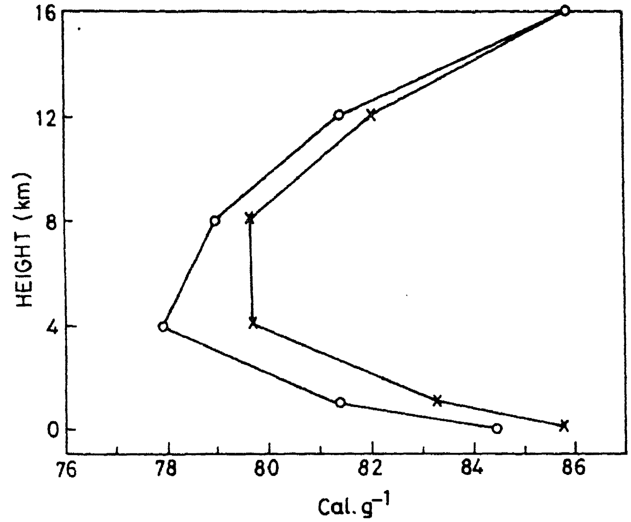


Fig 3 Vertical distributions of moist static energy (circle) and saturation moist static energy (cross)

Results

In this section results obtained from different experiments with SST values of 298K, 299K, 300K and 3.01K have been presented. The evolution of the circulation from the initial weak vortex in all the four experiments are compared to assess the model sensitivity and also to study the role of SST on the evolution of cyclonic circulation. In the following description experiments are referred to as Experiments I, II, III and IV to correspond with those of SST as 298K, 299K, 300K and 301K respectively.

The variation of the central surface pressure (CSP) with time in all the four experiments are presented in Fig. 4. In Experiment I the CSP reduced gradually till 24hr, then decreased steeply reaching a minimum of 981mb after 42hr period. In Experiment II, steep fall of CSP occurred after 18hr, reaching a minimum of 972mb after 48hrs. In Experiment III, the sharp fall of CSP started after 12hr reaching a minimum of 966mb after 42hr. In Experiment IV, the trend is similar as in the experiment with 300K. The first stage of gradual

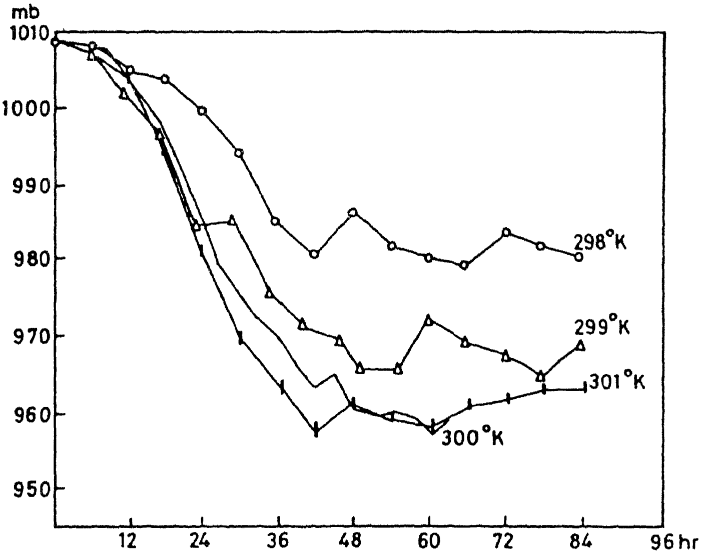


Fig 4 Time variation of central surface pressure (mb) for the four experiments with different SSTs

decrease of GSP corresponds to the pre-developing stage; the steep decrease of the CSP corresponds to the deepening stage and little variation of the CSP after reaching the minimum corresponds to the mature stage.

The time-radius sections of the tangential winds at 1km level for all the four experiments are shown in Fig. 5. In Experiment I (Fig. 5a) the distribution shows that the tangential winds reached a maximum of 33mps at 48hr at 60km radius. The core of the maximum is confined to the central 100km with the maximum occurring during 30hr to 70hr. In Experiment II (Fig. 5b), the tangential winds show a maximum of 39mps at 48hr and the maximum shows outward expansion till mature stage and inward slope later. In Experiment III (Fig. 5c) it is observed that the cyclonic winds reached an intensity of 42mps at 48hr at 100km radius. The tangential winds are noted to be stronger than the two previous experiments. In Experiment IV (Fig. 5d), the tangential winds reached a maximum of 40mps at 36hr. The maximum tangential wind moves inward towards the centre during the deepening stage and outwards later.

The structure of the circulation during the evolution of the vortex is presented corresponding to the mature stage only for all the four experiments. However, the results corresponding to the developing stage, though not presented, are referred to in the description to compare with the results. The height-radius sections of the tangential wind, radial wind, vertical velocity, potential temperature deviation and relative humidity are presented corresponding to the mature stage.

Experiment I (SST= 298K)

The features at the mature stage (*M*-stage) corresponding to this experiment are presented in Fig. 6. The tangential winds (Fig. 6a) show cyclonic circulation throughout the troposphere with a maximum of 30mps at 80km radi-

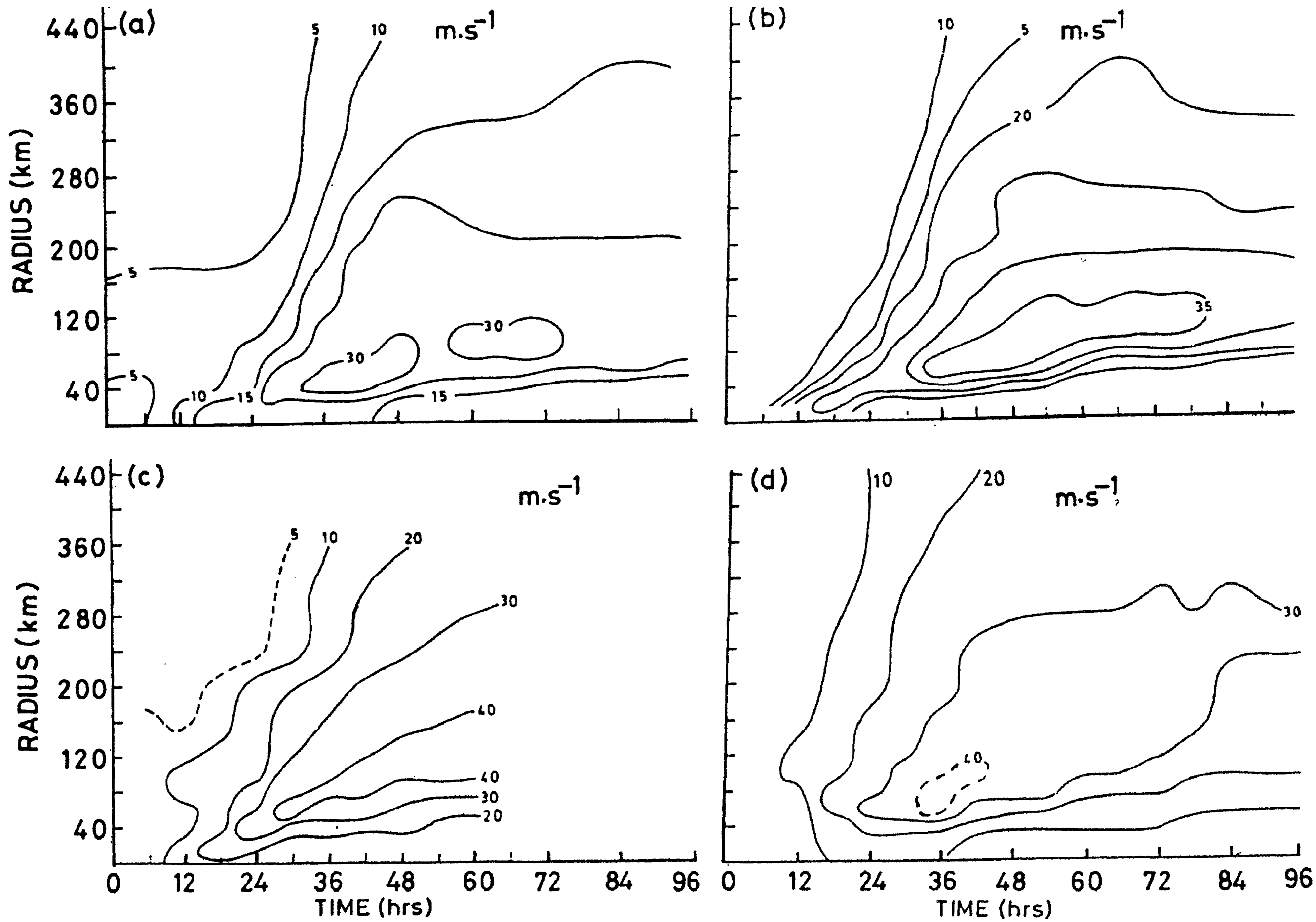


Fig 5 Time-radius section of tangential winds (mps) at 1km level for experiments with SST equal to (*a*) 298K; (*b*) 299K; (*c*) 300K and (*d*) 30

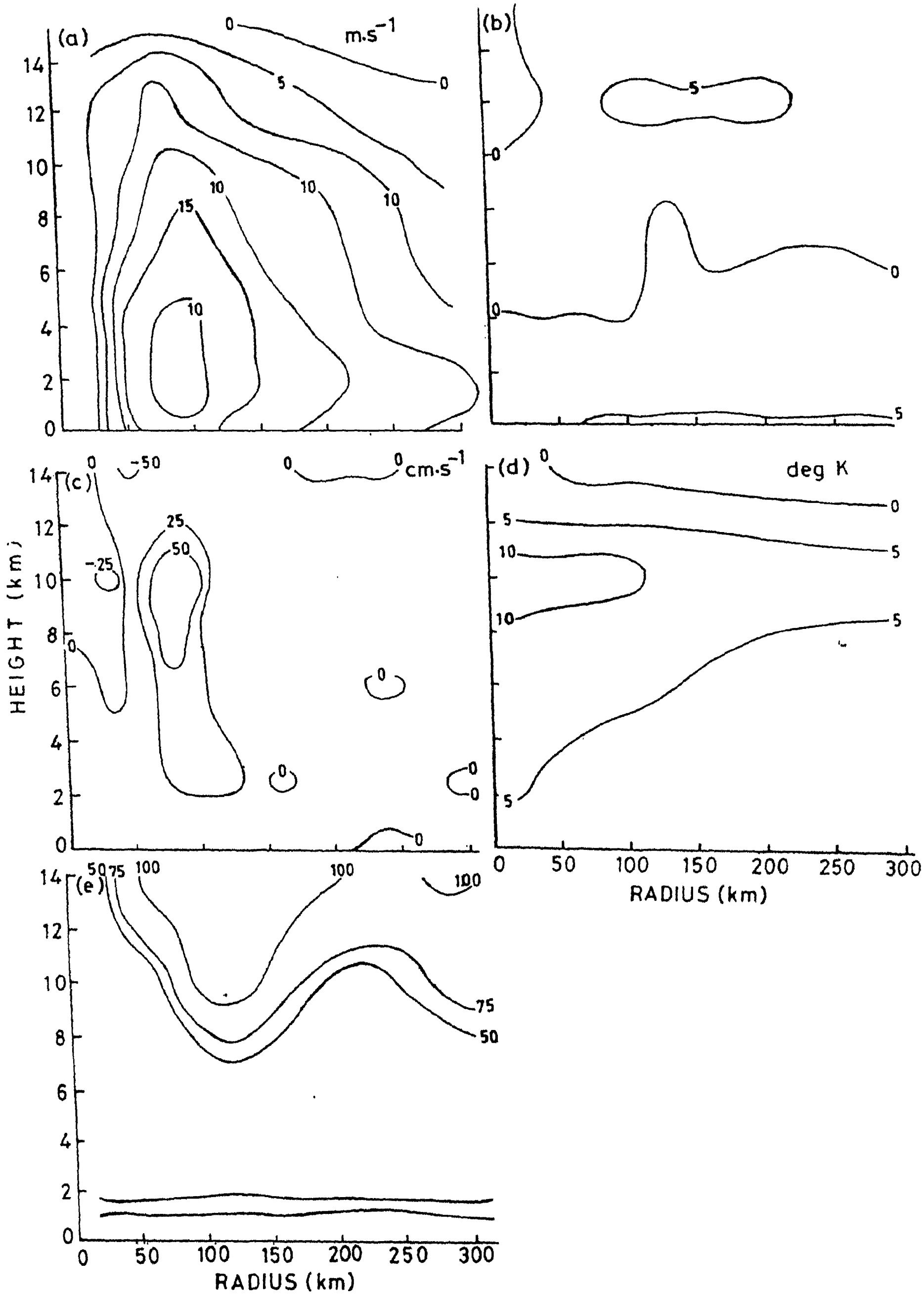


Fig 6 Height-radius distribution at mature stage of (*a*) Tangential wind (mps); (*b*) Radial wind (mps); (*c*) Vertical Velocity (cmps); (*d*) Potential temperature deviation (*K*) and (*e*) Relative humidity (%) for experiment – I i.e., SST = 298K

us in the lower troposphere. A weak anticyclonic circulation is found at very high levels. The radial wind field (Fig. 6b) shows inflow at lower levels and the outflow at higher leves, both with a strength of 5mps, showing a decrease from the developing stage. The vertical motion field (Fig. 6c) shows descending motion at the centre and vertical motion beyond 40km radius. The ascending motion which was very strong during the developing stage with intensities of 200cmps in respect of upward and downward motions decreased in strength of 50cmps after the mature stage. The strong ascending motion over a large region with subsidence at the centre indicates eye formation at the centre. The temperature field (Fig. 6d) shows heating throughout the troposphere with a maximum of 10K in the centre at the 8-10km height. The relative humidity field (Fig. 6e) shows drying at the centre throughout the core of the cyclone which corresponds to the subsidence. At higher levels moist region with 100% relative humidity is observed which coincides with the outflow region. However, there is a dry region in the lower troposphere with relative humidity less than 80%. This region of dryness in the region of convection is not realistic and is to be attributed to the cumulus induced subsidence effect of the Arakawa-Schubert scheme.

Experiment II (SST= 299 K)

The structure of the circulation at mature stage is presented in Fig. 7. The radius-height section of the tangential winds (Fig. 7a) shows that the cyclonic circulation is found to extend throughout the troposphere with a maximum of 39mps at 80km radius in the lower troposphere. A weak anticyclonic circulation is found at the uppermost level. The radial wind field (Fig. 7b) shows that both the outflow and inflow decreased from the developing stage. The maximum outflow region with 10mps is found at 12km height at a radial distance beyond 240km from the centre. The vertical motion field (Fig. 7c) shows strong descending motion at the centre with a maximum of 180cmps and a vertical motion beyond 40km with a maximum of 100cmps. This indicates the formation of eye. The temperature field (Fig. 7d) shows warming with the maximum reaching 15K at 10km height at a radial distance of 60km. The relative humidity field (Fig. 7e) shows drying in a narrow column throughout the core of the cyclone up to 11km which corresponds to the area of subsidence. At higher levels moist regions with relative humidity values of 100% are observed which can be attributed to the presence of an outflow region. In the lower troposphere, regions with less than 25% humidity are found and this unrealistic feature of dryness at lower levels, is also noted in previous experiment.

Experiment III (SST= 300 K)

The structure of the vortex at the mature stage i.e., at 48hr is shown in Fig. 8. The radius-height section of the tangential winds (Fig. 8a) shows that the tangential winds increased in strength reaching 43mps at 100km radius. This shows a shift of the maximum from 50km at developing stage to 100km at the mature stage. The cyclonic circulation extended throughout the vertical extent with anticyclonic winds occurring at 16km level. The radial wind field (Fig. 8b) clearly shows inflow at the lower levels and outflow at the upper levels. The strength of the outflow maximum decreased from 25mps at the deve-

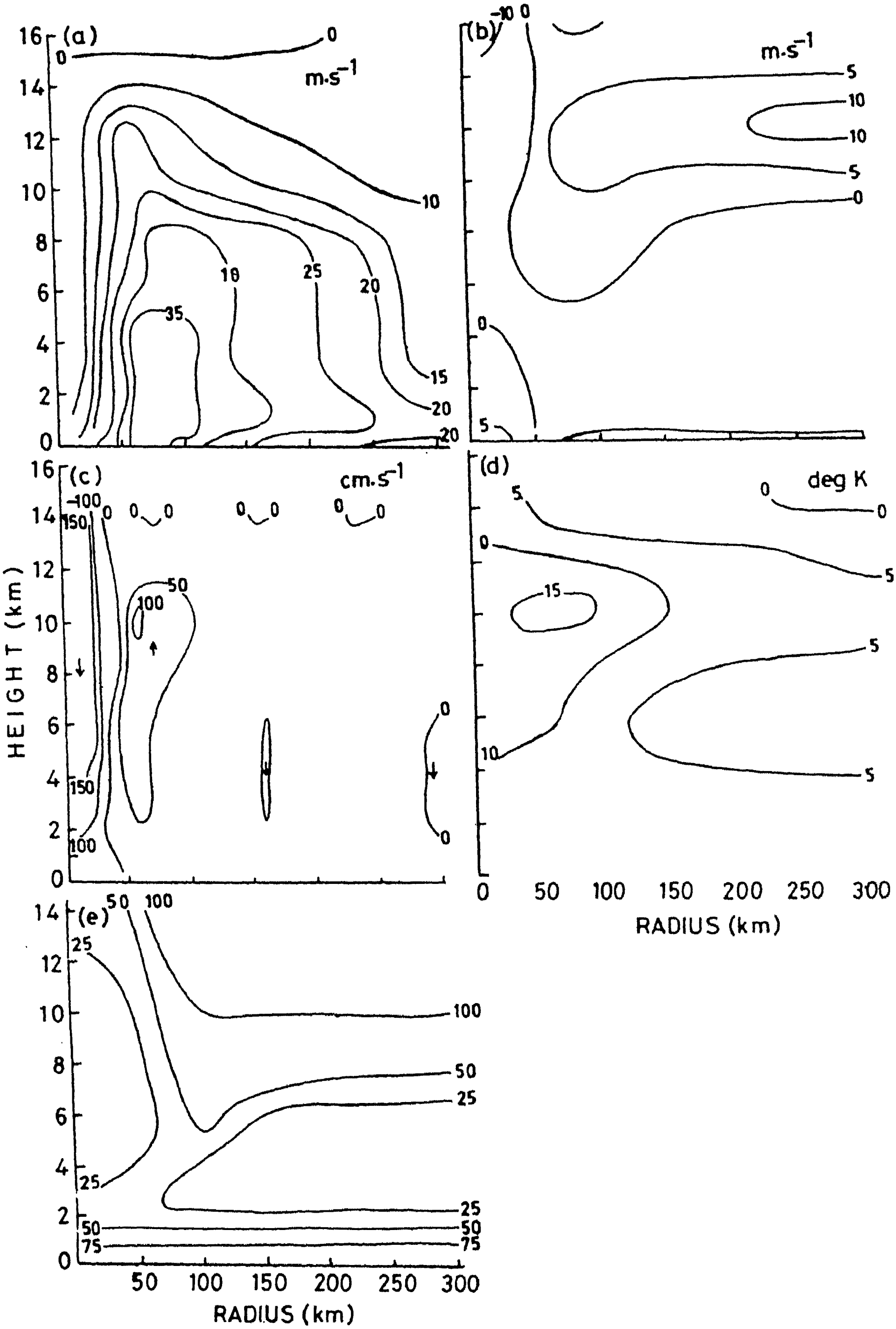


Fig 7 Same as Fig. 6 except for experiment – II i.e., SST = 299K

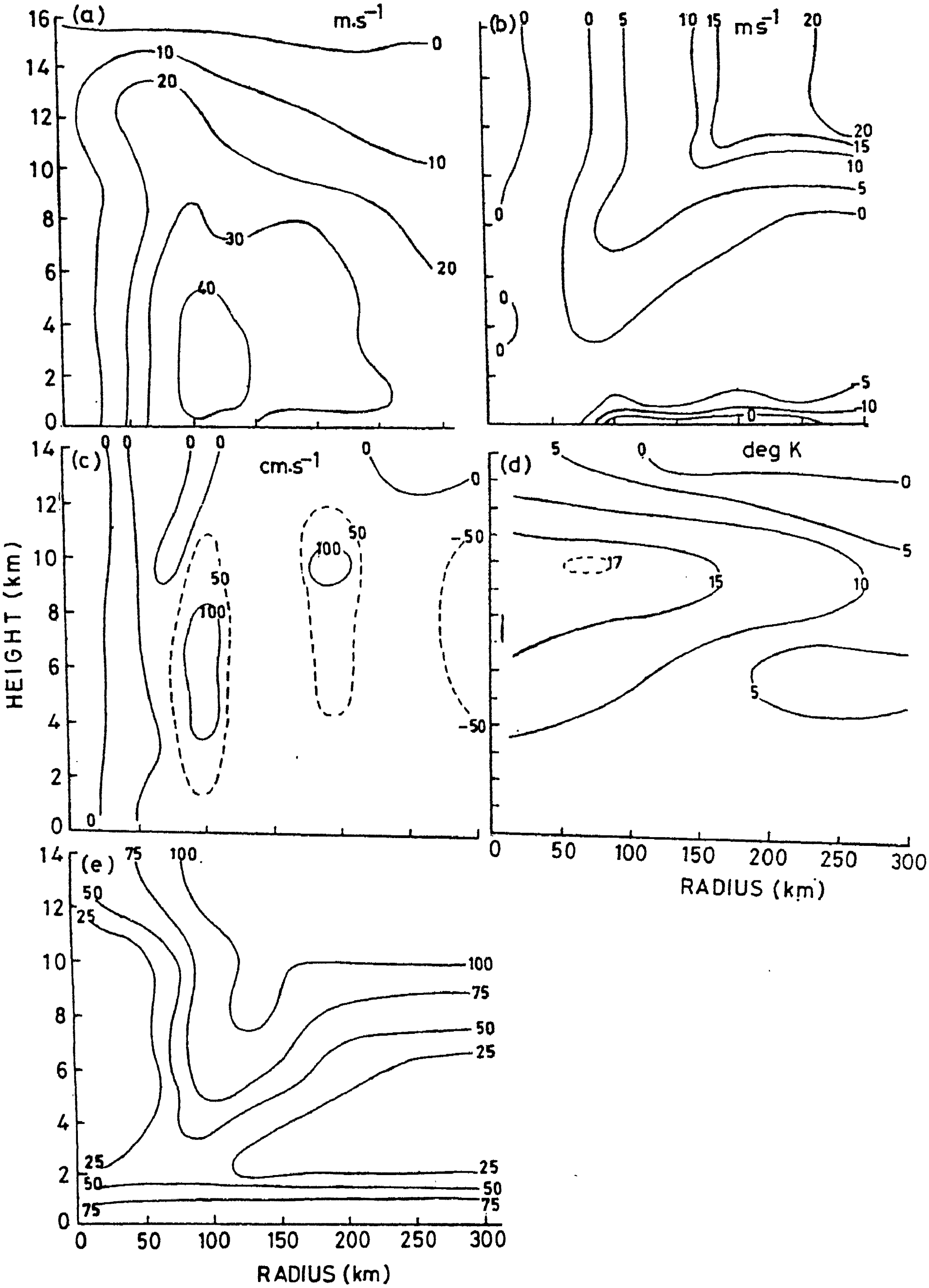


Fig 8 Same as Fig. 6 except for experiment – III i.e., SST = 300K

loping stage to 16mps at the mature stage. The inflow at the lower levels increased considerably from 12mps at the developing stage to 18mps at the mature stage. The vertical motion field (Fig. 8c) shows ascending motion beyond 40km radius and strong descending motion at the centre. This descending motion in conjunction with the absence of convection within the central 60km indicates the eye formation. The potential temperature deviation field (Fig. 8d) shows that maximum warming of 17K takes place at 10km level at a radial distance of 80km. The radial extent of warming also increased at the mature stage. The relative humidity field (Fig. 8e) shows more drying at the centre upto 60km. This is due to the subsidence at higher levels. A moist region which is the result of the outflow is also noted in this experiment and is similar to the one found in the previous experiments. A relatively dry region is found at the lower level of 2.5km is also a feature similar to the previous experiments.

Experiment IV (SST = 301 K)

The structure of the vortex at 48hr representing the mature stage is given in Fig. 9. It is observed that cyclonic circulation intensified with maximum tang-

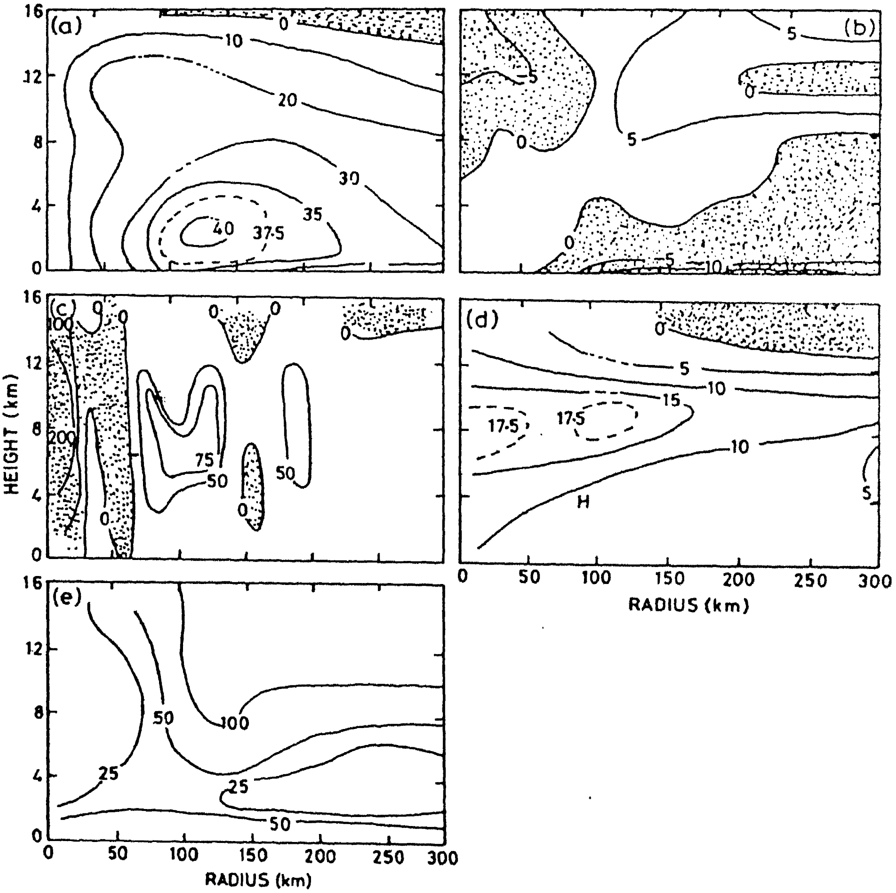


Fig 9 Same as Fig. 6 except for experiment – IV i.e., SST = 301K

ential winds of 40mps (Fig. 9a) observed between 100km and 150km radius indicating a shift from 50km radius at the developing stage. The circulation extended throughout the vertical extent with anticyclonic circulation appearing beyond 100km radius at the uppermost levels. The radial wind field (Fig. 9b) shows clear inflow and outflow regions. The strength of the outflow decreased to a maximum of 5mps from 20mps during the developing stage, while the inflow strength remained same. The vertical motion field (Fig. 9c) shows strong descending motion at the centre with an increase in strength up to a value of 200cmps at the centre. The ascending motion reduced in strength to 75cmps compared to the developing stage but two maxima are still noticed. This also corresponds to the reduction in convection especially in high clouds. The absence of convection at the centre and the strong descending motion clearly demonstrated the formation of the eye at the centre. The temperature field (Fig. 9d) shows further warming of the atmosphere with maximum warming taking place around the 8km level. The two maxima of almost equal magnitude show that the warming due to convection is as effective as the warming due to the strong subsidence at the core. The cold region has moved outward but still could be clearly noticed above the cyclonic circulation. The relative humidity field (Fig. 9e) shows strong drying in a narrow column at the centre within 50km radius corresponding to the region of subsidence. The moist region at higher levels associated with the outflow is also noted, but dry atmospheric condition with relative humidity of 25% is noticed at lower levels. This feature of dryness is similar to the previous experiments.

The time-radius sections of the cloud base flux for H-type clouds for all the four experiments are presented in Fig. 10. In the Experiment I (Fig. 10a) high clouds started from about 18hr corresponding with the developing stage. High clouds with maximum cloud mass flux have been observed from 40-200km after 24hr. In Experiment II (Fig. 10b), it is observed that the high type clouds formed from the beginning with convection values reaching 855gm/m^2/sec. at 24hr at a radius of 60km. The maximum convective region is extended during the mature stage and little convection is observed beyond 160km. In Experiment III (Fig. 10c), it is observed that the high type clouds started from 12hr with maximum cloud base mass flux of 400gm/m^2/sec. at 60km radius at 36hr. In Experiment IV, it is observed (Fig. 10d) that the high type clouds started to form from the initial time with strong convective.region at 100km radius. The maximum convection region shifted inward with time concentrating around 50-100km radius during the 24-36hr time period which corresponds to the developing stage.

Summary and Conclusions

The results described above are summarised as follows:

With the increase of SST the intensity of the simulated cyclonic circulation also increased. The minimum CSP attained varied from 981mb to 966mb when the SST increased from 298K to 301K, specifically when the SST increased from 298K to 299K the minimum CSP decreased by about 9mb and for the increase from 299K to 301K the decrease was 6mb. These results

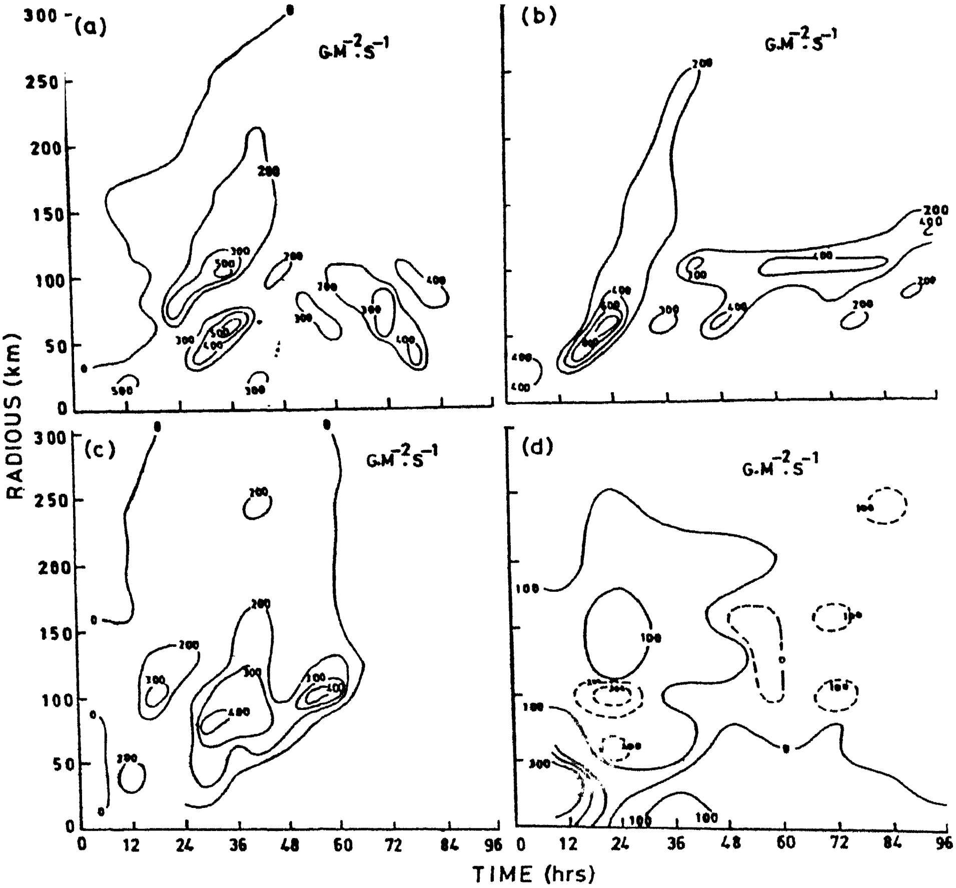


Fig 10 Time-radius section of cloud base mass flux (g m^{-2} s^{-1}) for high clouds for experiments with SSTs equal to (*a*) 298K; (*b*) 299K; (*c*) 300K and (*d*) 301K

agree with those of Sundqvist[17] and Tuleya & Kurihara[20] who found that the largest change in intensity is with SST equal to 300K. Correspondingly the tangential wind at 1km level showed an increase in the attained maximum of 33mps with SST at 298K; 39mps with SST at 299K and 40mps for SST at 300K and 301K. Apart from this the outward extent of the cyclonic circulation has also increased with increase of SST. When the SST has increased from 300K to 301K, the maximum tangential wind has been attained earlier by 18hrs. The evolution of the circulation has shown similar features in all experiments. The important features of the evolution such as the development of the cyclonic circulation, inflow and outflow fields, correspondence with the ascending the descending motions and warm core structure are similar but with variations in strength. However, it is observed that the convective features have been better organized when SST has been increased from 298K to 299K. It is also noticed that the core of the system is warmer with the increase of SST indicating that the convection has increased with the increase of SST.

From the above observed features, it may be inferred that though the broad features of the evolution are similar, it is clear that higher SST values are favourable for the disturbances to grow stronger. It is also noticed that the development of the system has conspicuously changed when the SST has been increased from 298K to 299K indicating that SST values less than 25.8°C are unfavourable for the development of the disturbance into an intense cyclonic system. From the observation that there are little variations in the intensity of the circulation when the SST has increased 300K to 301K it may be inferred that SST values higher than 300K may not be effective in producing stronger cyclonic systems but only be effective for a faster development.

The above experiments indicate that the model used is capable of producing the features of the evolution of cyclonic circulation from a weak vortex. The model is capable to produce changes in the cyclone intensity in relation to SST variation. An important result is that the model derived circulation shows that a conspicuous change when SST increased from 298K to 299K thus upholding that a minimum SST of 26°C is necessary for cyclogenesis. It may also be necessary to use a model with better resolution and to experiment with smaller SST intervals for a better understanding of the associated changes in the cyclone evolution.

Acknowledgement

The authors thank Mr K Ashok for his useful discussions and for his kind help in the preparation of the diagrams and the manuscript. This work has been carried out with the support of DST for a scientific project under grant No. 12(14)84/STP/II dt 8-11-1985.

References

1 A Kashara *J Meteor* **18** (1961) 259-282
2 S Syono *Proc int Symp Num Wea Pred Tokyo* (1962) 405-418
3 J G Charney and A Eliassen *J atmos Sci* **21** (1964) 68-75
4 H L Kuo *J atmos Sci* **22** (1965) 40-63
5 H L Kuo *J atmos Sci* **31** (1974) 1232-1240
6 A Arakawa and W H Schubert *J atmos Sci* **31** (1974) 674-701
7 M Yamasaki *Geophys* **19** (1968) 559-585
8 S L Rosenthal *ESSA Tech Mem ERLTM-NHRL* **82** US Dept. of Commerce, National Hurricane Research Laboratory Miami Fla, USA (1969) 36pp
9 H Sundqvist *Tellus* **22** (1970) 359-390
10 R A Anthes *Mon Wea Rev* **99** (1971) 617-635
11 M Wada *Met Met Soc Japan* **57** (1979) 505-530
12 D V Bhaskar Rao 11 *Nuov Clim* **10C** (1987) 677-696
13 R A Anthes *Meteorol Monogr* **19**(41) (1982) American Meteorological Society 209 pp
14 E Palmen *Geophysica* **3** (1948) 26-38
15 W M Wendland *J appl Met* **16** (1977) 477-481
16 K Ooyama *J atmos Sci* **26** (1969) 3-40
17 H Sundqvist *Tellus* **22** (1970) 504-510
18 S W Chang *Mon Wea Rev* **110** (1979) 2063-2069
19 S W Chang and R V Madala *J atmos Sci* **37** (1980) 2617-2630
20 R E Tuleya and Y Kurihara *Mon Wea Rev* **110** (1982) 2063-2069
21 S J Lord, W C Chao and A Arakawa *J atmos Sci* **39** (1982) 104-113
22 C L Jordan *J Meteor* **15** (1958) 91-97

A REAL TIME STORM SURGE PREDICTION SYSTEM: AN APPLICATION TO EAST COAST OF INDIA

S K DUBE, A D RAO, P C SINHA AND P CHITTIBABU
Centre for Atmospheric Sciences, Indian Institute of Technology, Hauz Khas, New Delhi 110 016, *India*

(*Received* 20 *September* 1993; *Accepted* 23 *September* 1993)

The east coast of India is threatened by the possibility of storm surge floods whenever a tropical cyclone approaches. Storm surge disasters cause heavy loss of life and property, damage to the coastal structures and losses of agriculture which leads to annual economic losses in India. Thus the real time monitoring and warning of a storm surge is of great concern to India. The goal of this paper is to describe a real time storm surge prediction system which has been applied to the east coast of India..

The storm surge prediction system is based upon a numerical model developed earlier[1,2]. Surface winds associated with a tropical cyclone are derived from a storm model[3] which requires only the positionsof the storm, the pressure drop and the radii of maximum winds. The model has been used to simulate the surge generated by severe tropical cyclones during last three decades. The model computed surge is found tobe in good agreement with available observations.

Key Words: Storm Surges; Bay of Bengal; Tropical Cyclone; Numerical Model

Introduction

Storm surges that are associated with severe tropical cyclones constitute one of the world's foremost natural disasters. Amongst other natural disasters, the storm surge stands out as the most damaging and, indeed, as an agent of death and destruction on a scale as massive as that of earthquakes. Death and destruction arise directly from the intense winds characteristic of tropical cyclones blowing over a large surface of water, which is bounded by a shallow basin. As a result of these winds, the massive piling up of the sea water occurs at the coast leading to the sudden inundation and flooding of coastal regions.

In India, the study of the numerical storm surge prediction was pioneered by Das[4]. He conducted a numerical experiment and computed the surge generated by an idealised cyclone striking the coast of Bangladesh. The above study was extended for application to the coast of West Bengal and North Orissa[5]. An objective method to predict the storm surges on the entire east coast of India north of 10°N was also evolved[6]. This consists of precomputed nomograms based on the output of a SPLASH model[7] developed in the United States. Nonlinear models for the simulation of surges along the Bangladesh coast in which the dynamic effects of the Ganga-Brahmaputra-Meghna river system; as well as the effects of islands near the Meghna estuary have been incorporated, were also developed[8,9].

In these storm surge predictions for the Bay of Bengal, the coastal boundary is approximated by conventional orthogonal straight line segments. An effective means of achieving greater accuracy is a precise representation of the shoreline which is curved and/or broken by bays or estuaries. In storm surge models, curvilinear boundary representation depending on the application of conformal mapping have been previously considered by various workers for other regions of the world[10]. A numerical storm surge prediction model for the Bay of Bengal was also developed which is capable of dealing with curved coasts[11]. The model is fully nonlinear and has been used to simulate the surges generated by several cyclonic storms[1,2].

Operational Models

Operational numerical storm surge prediction models have been developed and are being routinely used for several coastal regions of the world such as the North Sea[12,13], the Gulf of Mexico, the Atlantic coast[14], Hong Kong and China. A review of these models is given in the several publications[15-18].

Most models require large computing power. Therefore, for routine prediction of storm surges, especially for providing multiple forecast scenarios, the models cannot be used without access to sufficient computing facility. To overcome this difficulty, most of the forecasting offices, including the India Meteorological Department, use nomograms for prediction of storm surges associated with tropical cyclones[7]. The nomograms have been developed from modelling studies of a large number of bathymetries and approach angles.

The second WMO International Workshop on Tropical Cyclones[19] recommended the use of personal computers (PCs) by developing countries to adopt a stand-alone storm-surge forecasting system. The recent advent of powerful personal computers has opened up the possibility of running dynamical models in real time on PC-based work stations in an operational office. In fact, a PC-based work station (the Automated Tropical Cyclone Forecasting System, ATCF) is already in operation at the Joint Typhoon Warning Centre, Guam for many years. The Australian Bureau of Meteorology Research Centre, together with their Bureau of Severe Weather Programme Office has also developed an Australian Work Station for storm surge forecasting. The purpose of the present paper is to describe a real time storm surge prediction system for the east coast of India.

Basic Features of the Model

In this section we will briefly describe the basic features of the model proposed for real time storm surge prediction.

Dynamic Storm Model

In the proposed predictive system the surge is generated by a cyclone, moving across the analysis area. In view of the strong associated winds and high values of wind stress, the forcing due to barometric changes have been neglected. Thus, the surface wind field associated with a tropical cyclone is the primary requirement for modelling storm surges. The wind field at the sea surface is derived by using a dynamic storm model[3]. This model uses as input the radius of maximum wind and the pressure drop. The main portion of the storm model is a trajectory model and a wind speed profile approximation scheme. The trajectory model represents a balance between pressure gradient, centrifugal, coriolis and surface frictional forces for a stationary storm.

Storm Surge Model

A vertically integrated numerical surge prediction model developed earlier was adopted for surge prediction. This may be used as a menu-driven standalone system. The details of the model and the numerical solution procedure are described in earlier publications[1,2]. Only a brief description of specific features will be presented here.

The model is nonlinear and is forced by the wind stress and by quadratic bottom friction. We find that the nonlinear advection terms have significant effect on the final results, especially in the shallow coastal waters of the head Bay of Bengal. Therefore, for operational applications the nonlinear terms cannot be left out. The treatment of the coastal boundaries in the model involve a procedure leading to a realistic curvilinear representation of both the western and the eastern sides of the Bay of Bengal. This coastal representation has an added advantage of taking automatically into account the finer resolutions in the shallow regions of the northern Bay. A desirable feature of storm surge simulation scheme is the ability to incorporate increased resolution adjacent to the coastline. This has been achieved by using a variable grid which leads to substantial refinement of resolution near the coastline, and a coarser resolution in deeper waters. This type of grid assists in the incorporation of a more detailed bathymetric specification in the important coastal region. With the specification of east-west grid points, the first off-shore grid points at which the elevation is computed is, on an average about 5km from the coastline.

Integration Procedure

A conditionally stable semi-explicit finite difference scheme with a staggered grid is used for the numerical solution of the model equations[20]. The staggered grid consists of three distinct types of computational points on which the sea surface elevations and the zonal and meridional components of depth-averaged currents are computed. The computational stability is achieved by satisfying the CFL (Courant-Friedrich-Lewy) criterion. In the present model this condition is satisfied by limiting the time step of integration to 3 minutes.

Bottom Stress

The bottom stress is computed from the depth integrated currents using a conventional quadratic law with a constant coefficient of 0.0026. This value of the bottom friction coefficient has been decided by performing several numerical experiments[21].

Boundary and Initial Conditions

The coastal boundaries are taken as vertical side walls across which the normal transport vanishes. The normal currents across the open sea boundaries are prescribed by a radiation type of condition[22]. As usual it is assumed that the motion in the sea is generated from an initial state of rest.

Bathymetry

The bathymetry for the model is derived from the Naval Hydrographic Charts and is interpolated at the model grid points by using cubic spline interpolation scheme. With this procedure sufficiently accurate and realistic bathymetry is generated over the continental shelf.

Operating Procedure

The storm surge prediction model is run from a terminal menu. The steps are as follows:-

a. First the user sets up the forecast domain by executing an appropriate window. The user also provides an arbitrary number of stations around the forecasted place of landfall of the cyclone for peak surge display. The duration of the forecast or the number of iterations are also provided.

b. The user next provides the required characteristics of tropical cyclones for computing the wind stress forcings. These include: (*i*) Cyclone positions (Lat.-Long.), (*ii*) Pressure Drop (hpa), and (*iii*) Radii of maximum winds (meters) at any time interval (preferably six hourly observations). The input data are determined from INSAT imageries, cyclone detecting radars and a surface synoptic analysis by the India Meteorological Department.

c. The final step is to run the storm model and the surge model. The storm model produces the required surface wind stress associated with the tropical cyclone at the model time steps. The output of the surge model are sea-levels, current fields and the peak surge envelope for three forecast scenarios: (*i*) cyclone landfall at the forecast location on the coast, (*ii*) landfall at some distance (preferably 50 or 100km) to the left of the forecast location, and (*iii*) landfall to the right of the forecast location.

d. These fields may be displayed in any order or combination from the menu. The user may zoom into any part of the coastal region to get a clear picture if needed.

e. The whole process of running the model for a 48 hour forecast takes 15-16 minutes of time.

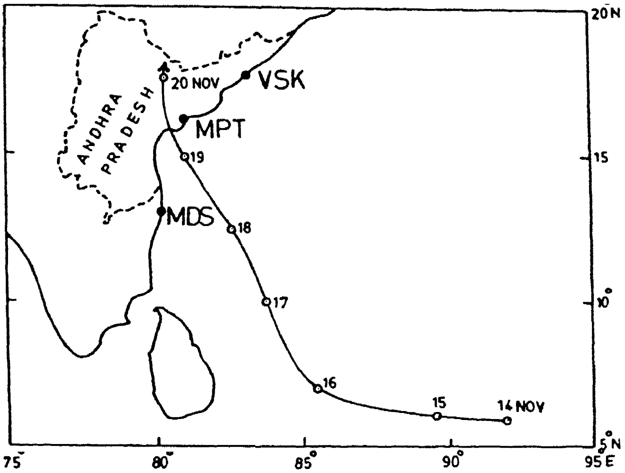
One of the significant features of this storm surge forecast system is its ability to investigate multiple forecast scenarios to be made in real time. For example, three 48-hour forecasts can be carried out on a 25 MHz PC-AT 386 in around 15 minutes. As the cyclonic storm approaches the coast, and the Meteorological Department's forecast become more accurate, the track of the cyclone is updated.

Validation Experiments

To validate the model, several simulation experiments have been performed by using the data of severe cyclonic storms hitting the east coast of India. In the present paper, an attempt was made to compare the simulated sea surface elevations with observations from local tide gauges where ever possible, or with post storm survey estimates of the India Meteorological Department.

November 1977 *Andhra Cyclone*

A severe cyclonic storm with a core of hurricane winds hit the Andhra Coast of India on 19th November 1977. The genesis of the cyclone was in the Malaysia region from where it moved westward as a low pressure area. It concentrated into a deep depression with its centre near 6°N, 92°E on the morning of 14th November. The subsequent history of the storm is shown in Fig. 1.



Fig·1 Track of Andhra Cyclone: 14-20 November, 1977. ⊙ Position at 0300 UTC

Whilst continuing to move generally westwards, the system rapidly intensified and on the morning of 16th, lay near 7°N, 85.5°E with a well-developed core of winds exceeding 30ms^{-1}. On the evening of 16th November, the storm changed its course towards a more northerly track, and continued to follow this until landfall on the Andhra coast. Satellite reports indicated that the storm attained its peak intensity on the 18th, and maintained this intensity on the 19th. The maximum wind speed associated with the storm during this period was estimated to be about 70ms^{-1} at a distance from the centre of about 40km. A pressure drop of 80hpa was estimated at the centre of the cyclone.

A major part of the destructive effect of the cyclone was felt along the main coast south of Masulipatnam in Andhra Pradesh. This resulted from the surge-induced sea-surface elevation which caused widespread flooding in the low-lying region, claiming a heavy toll of human lives. Sea-surface elevations of 5 meters above mean-sea level were estimated at Divi Island on the basis of floating debris which had become entangled in trees.

A two day simulation of the surge generated by the cyclone was commenced at 1200 UTC on November 17. Model computed peak surge values associated with the storm are shown in Fig. 2. It may be seen that the maximum simulated surge of 5 meters occurs at a grid point close to Divi Island. This is in very good agreement with the post storm survey estimated sea level elevation of about 5 meters in that region.

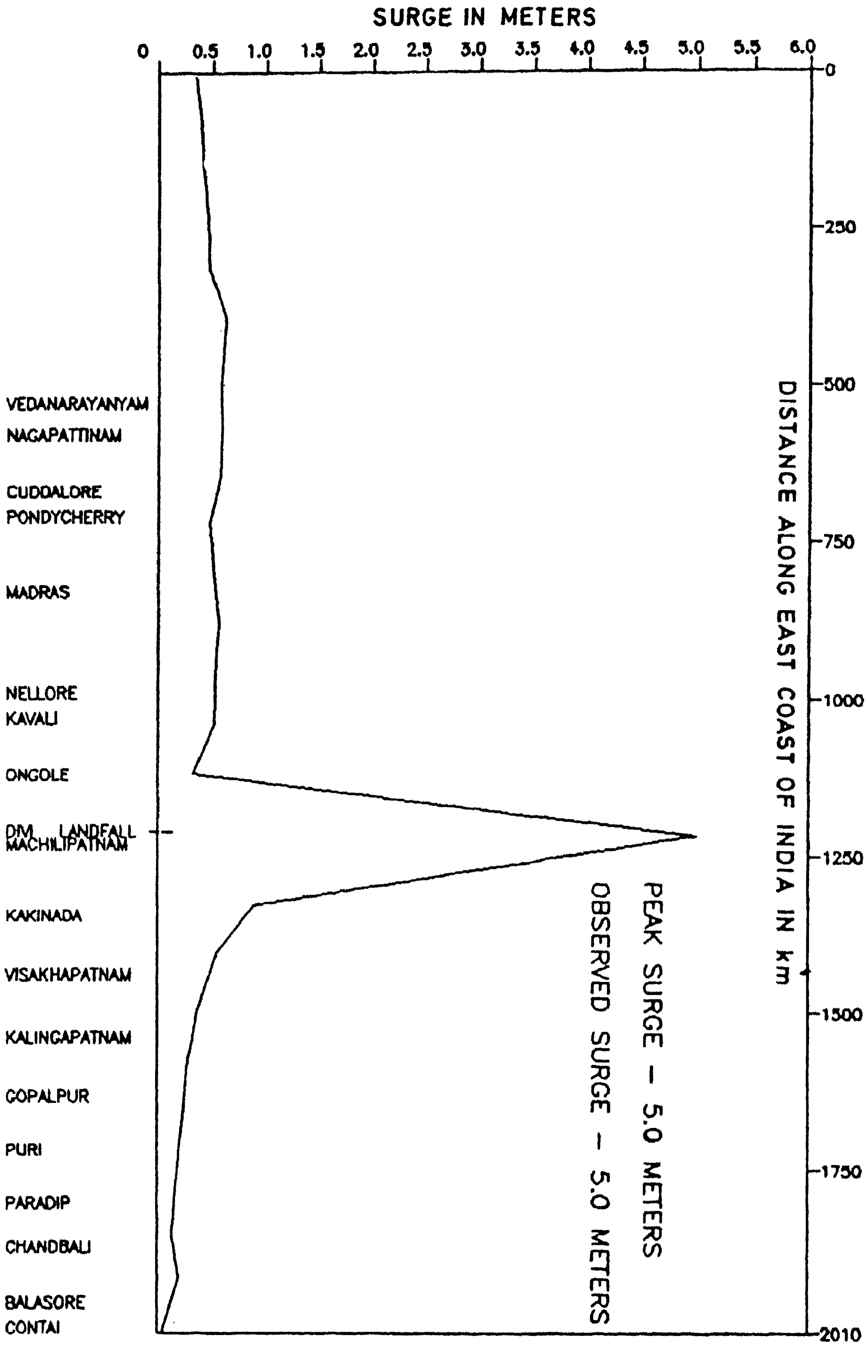


Fig 2 Peak Surge Envelope Associated with 1977 Andhra Cyclone

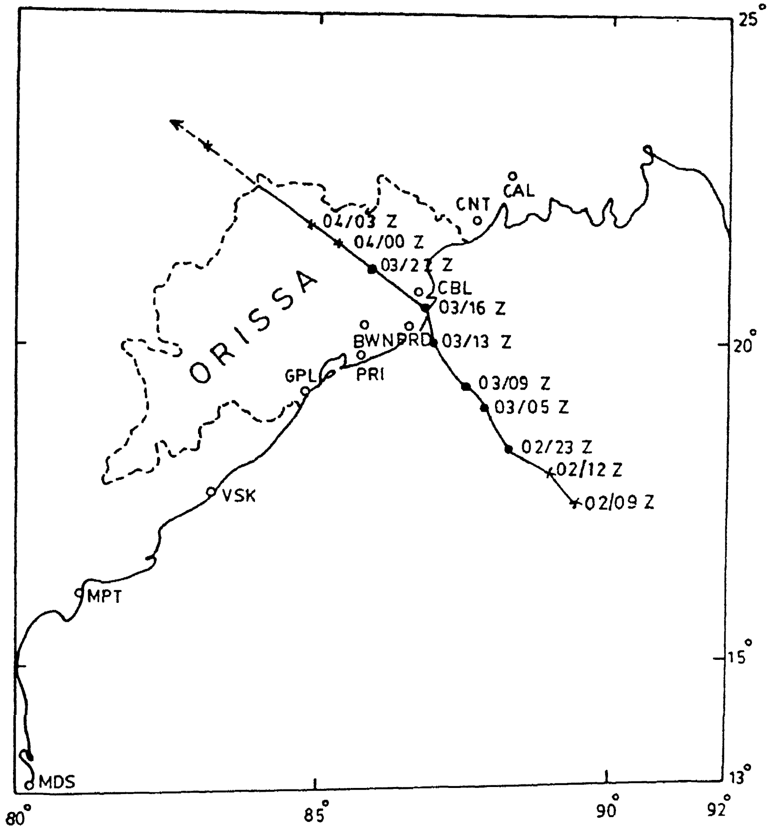


Fig 3 Track of Orissa Cyclone: 2-4 June, 1982. X Synoptic fixes of Storm Centre, ● Radar fixes of Storm Centre

June 1982 *Orissa Cyclone*

A cyclonic storm which formed over the east central Bay of Bengal on 1st June 1982 and was expected to cause a gale wind speed of 20 to $25ms^{-1}$ suddenly intensified into a hurricane with a wind speed of 45 to $60ms^{-1}$ in the afternoon of 3rd June. The storm crossed the North Orissa coast between Paradeep and Chandbali at 1800 UTC on the same day with a wind speed of $60ms^{-1}$ (Fig. 3). The Meteorological Department estimated a pressure drop of 50hpa and a radius of maximum winds of 40km at landfall (from satellite, radar and surface observations). Observations of peak sea-surface elevations were estimated to be more than 3 meters at Dhamra port which is about 55km to the right of the landfall. This is in close agreement with the model computed peak surge (Fig. 4).

May 1990 *Andhra Cyclone*

A severe cyclonic storm with a core of hurricane winds crossed the Andhra coast at the mouth of river Krishna on the evening of 9th May 1990 at

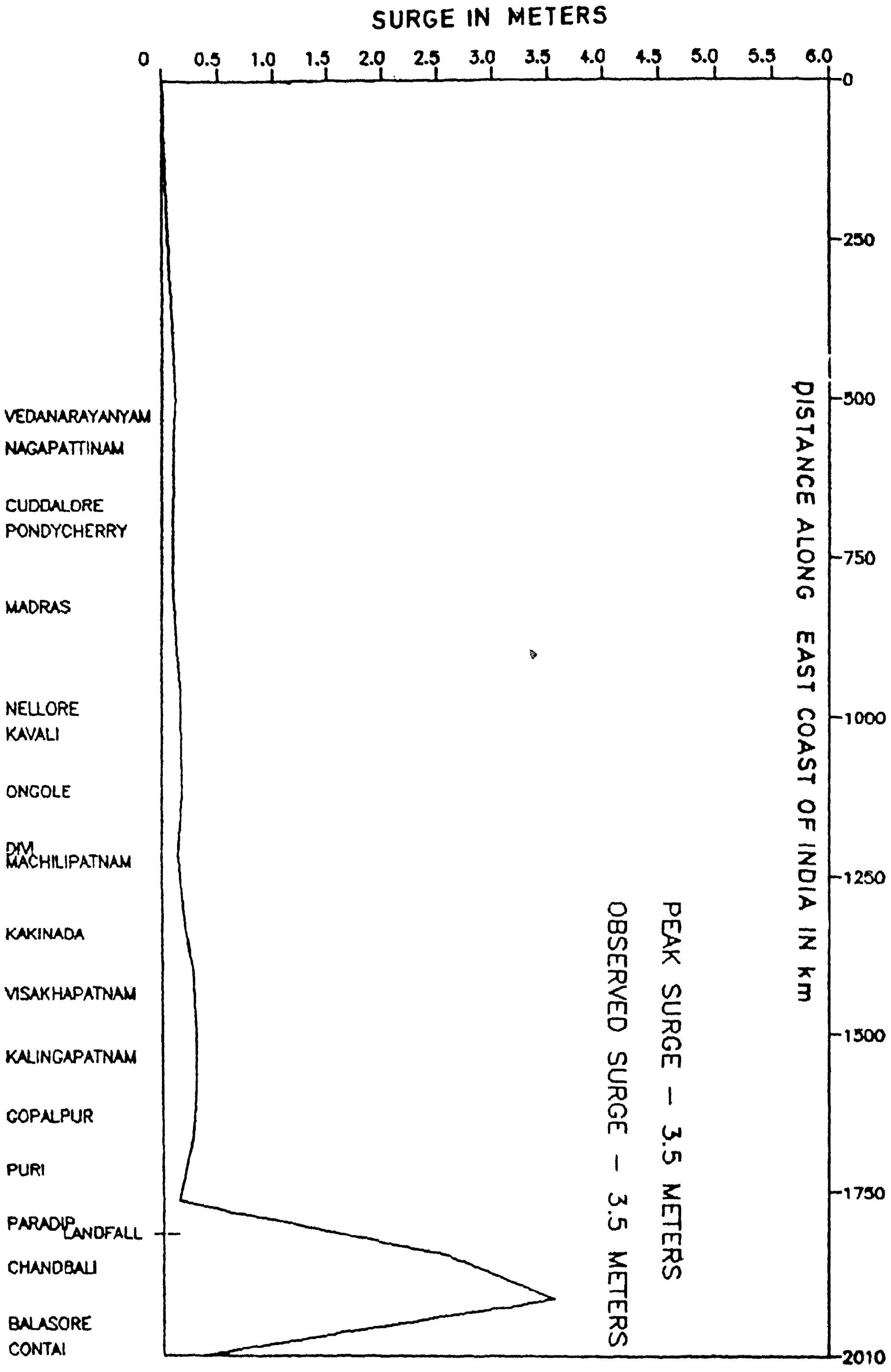


Fig 4 Peak Surge Envelope Associated with 1982 Orissa Cyclone

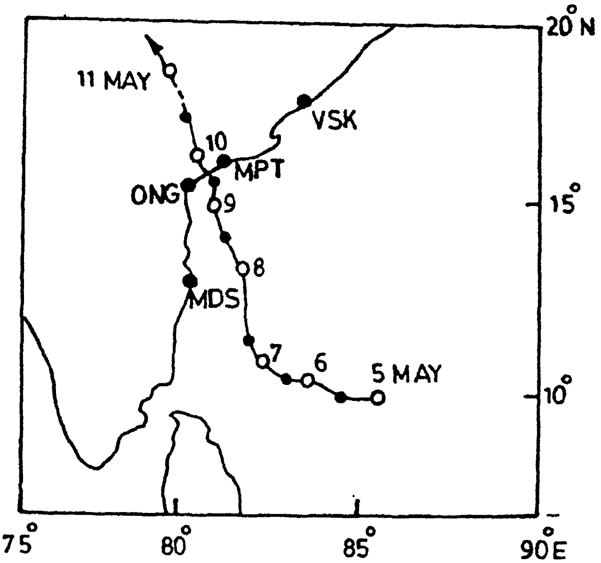


Fig 5 Track of Andhra Cyclone: 5-11 May, 1990. ○ Position at 0300 UTC and ● Position at 1200 UTC

about 1300 UTC (Fig. 5). The cyclone started as a well marked low pressure area over southeast Bay of Bengal and rapidly concentrated into a deep depression at 0300 UTC of 5th May, 1990 with its centre near 10.0°N and 85.5°E. On that evening it intensified into a cyclonic storm and was centered near 10.0°N and 84.5°E. Moving in a west-northwesterly direction it intensified further into a severe cyclonic storm by 6th morning when it was centered near 10.5°N, 83.5°E. The cyclone attained hurricane intensity around 0600 UTC of 6th May. It moved in west-northwesterly direction till 7th morning and changed its course to a north-northwesterly direction till its landfall on 9th May.

Based on the observation of INSAT-1B, the Meteorological Department estimated the lowest central pressure of the cyclone to be 920hpa and the maximum sustained winds of 127 knots at about 25km from the centre.

The cyclone caused heavy damage to life and property in the coastal region. A surge of 4-5 meters was reported from the region. At Edirumondi the level of Krishna river rose by about 4 meters and at Nagayalanka, about 25km north of Edirumondi, the rise in level was reported as 3-4 meters. Figs. 6 and 7 depict the model computed peak surge envelop and a 3-dimensional overview of the surge along the east coast of India. It may be noted that the model computed result of maximum surge height compares favourably with observed estimates. Figs. 8 and 9 provide two other forecast scenarios when the landfall of the cyclone is 100km to the south and north respectively of actual landfall point.

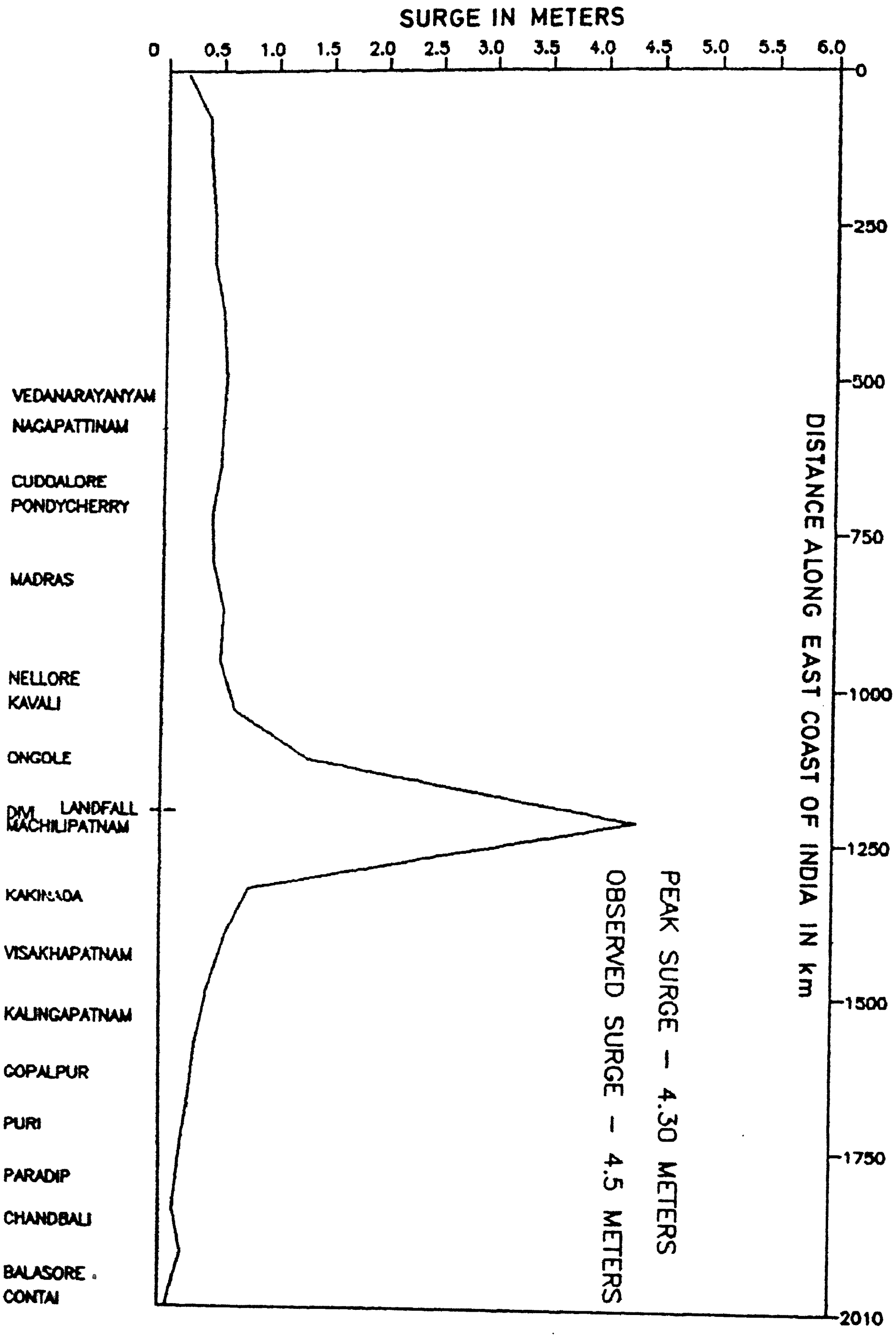


Fig 6 Peak Surge Envelope Associated with 1990 Andhra Cyclone

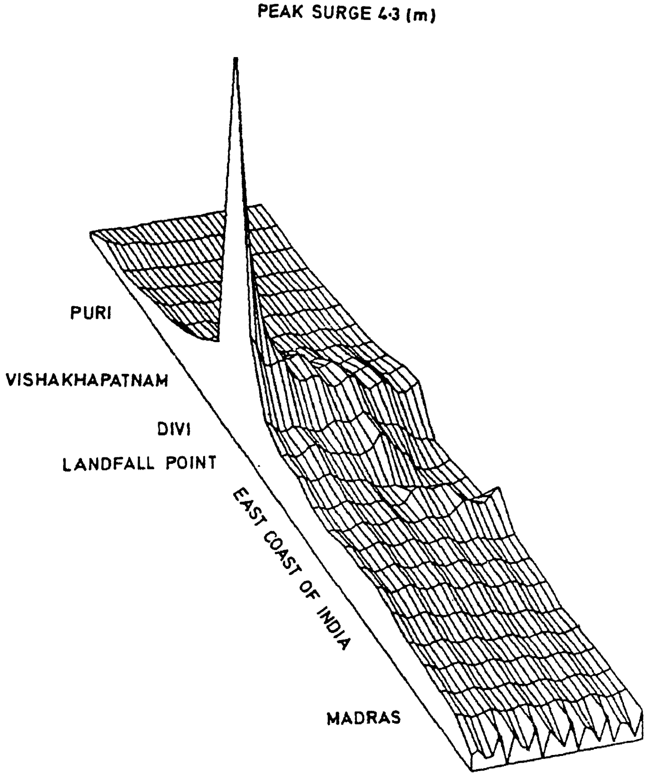


Fig 7 3-D Overview of the Peak Surge Envelope Associated with 1990 Andhra Cyclone

Conclusions

The development of a storm surge forecasting system has been described. The model can be run in a few minutes on a personal computer. The forecasting system is based on the vertically integrated numerical storm surge model[1,2]. A dynamic storm model[3] is used for computation of surface winds associated with cyclonic storms. The only meteorological inputs required for the model are the positions of the cyclone, pressure drop and radii of maximum winds at any fixed interval of times. The system is operated *via* a terminal menu and the output consists of the 2-D and 3-D views of peak sea surface elevations with the facility of zooming the region of interest. The system can handle multiple forecast scenarios.

The model results reported for three case studies in the paper are in very good agreement with the available observations and estimates of the surge. The model has extensively been tested with severe cyclonic storms of the last three decades which have affected the Indian coastal region.

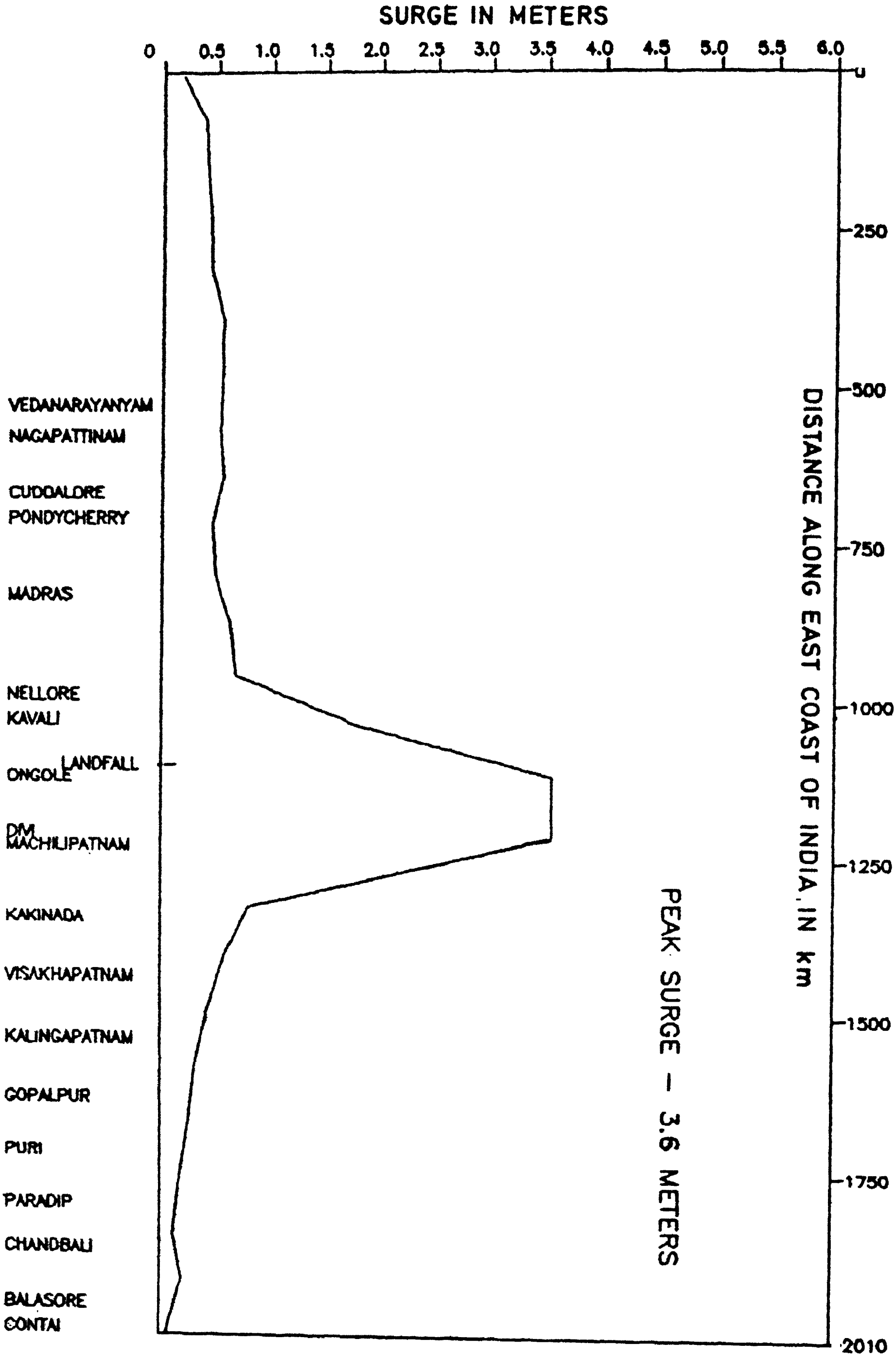


Fig 8 Peak Surge Envelope-Scenario I: When the landfall is 100km to the South

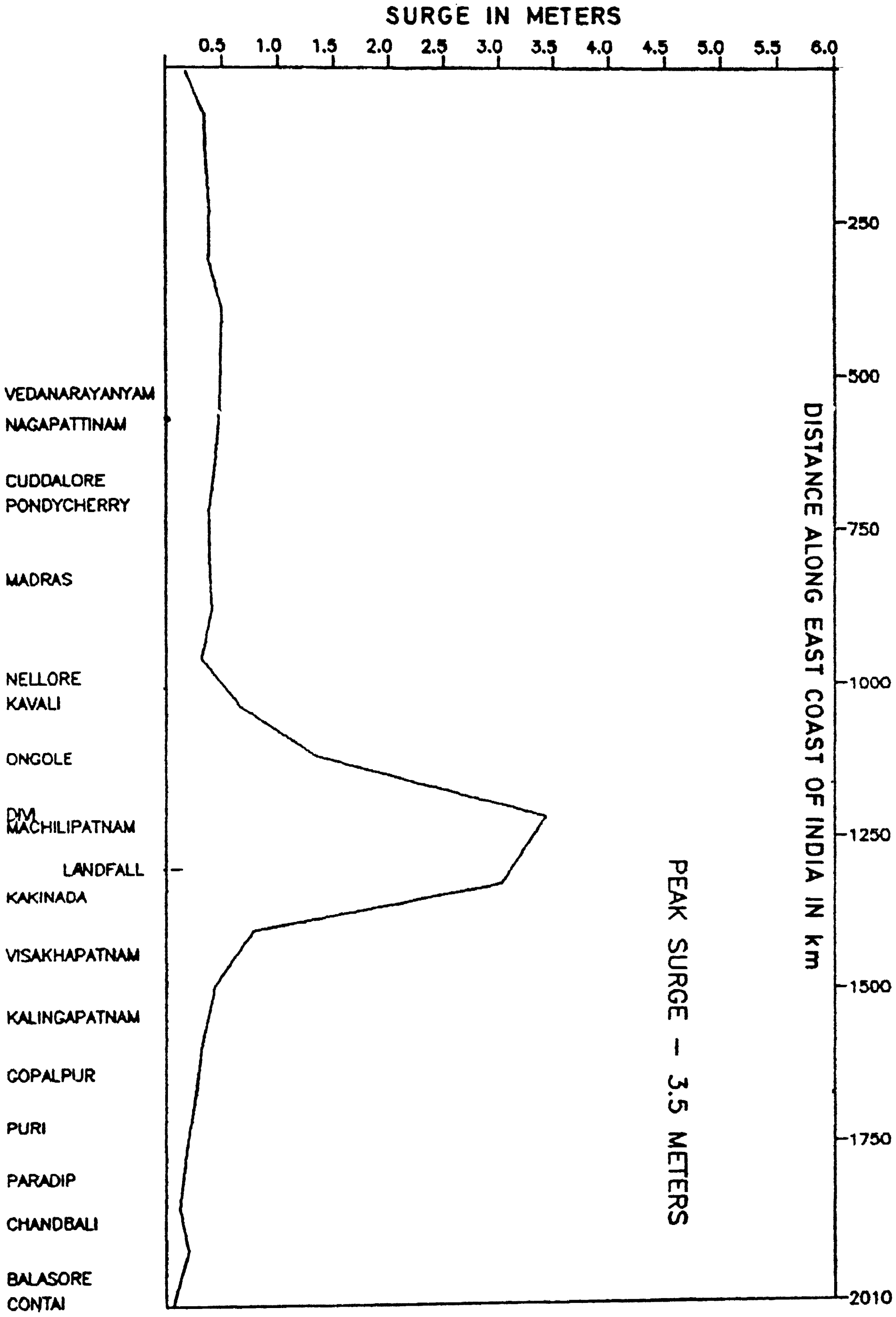


Fig 9 Peak Surge Envelope-Scenario II: When the landfall is 100km to the North

Acknowledgements

The authors are thankful to the Department of Ocean Development, Govt of India for supporting this research at Centre for Atmospheric Sciences, IIT, Delhi under its Coastal Ocean Design and Prediction System (CODAPS) Programme.

References

1 P K Das, S K Dube, U C Mohanty, P C Sinha and A D Rao *Mausam* **34** (1983) 359-366

2 S K Dube, P C Sinha and G D Roy *Dyn Atmos Oceans* **9** (1985) 121-133

3 C P Jelesnianski and A D Taylor *NOAA Technical Memorandum* ERL, **WMPO-3** (1973) 33pp

4 P K Das *Nature* **239** (1972) 211-213

5 P K Das, M C Sinha and V Balasubramanyam *Q J R met Soc* **100** (1974) 437-449

6 S K Ghosh *Indian J Met Geophys* **28** (1977) 157-168

7 C P Jelesnianski *SPLASH I: Landfall Storms, NOAA Tech Memo* NWS TDL-46 (1972) Washington DC, 52pp

8 B Johns In *Monsoon Dynamics* (*Eds* M J Lighthill and R P Pearce) Cambridge Univ Press (1981) 690-705

9 B Johns and A Ali *Q J R met Soc* **106** (1980) 1-18

10 C P Jelesnianski *Mon Wea Rev* **93** (1965) 343-358

11 B Johns, S K Dube, U C Mohanty and P C Sinha *Q J R met Soc* **107** (1981) 915-934

12 R A Flather *Rep Inst Ocean Sci No.* 30 (1976) Bidston UK

13 R A Flather and R Proctor in: *North Sea dynamics* (*Eds.* Sündermann/Lenz) Springer-Verlag Berlin Heidelberg (1983) 299-317

14 C P Jelesnianski and J Chen *Rep TDL, NWS* Silver Spring MD (1979) 16pp

15 C Jiping, T S Murthy, B Chenglan,.M I El-Sabh and L Fengshu China Ocean Press (1990) 326pp

16 C P Jelesnianski Second International Workshop on Tropical Cyclones, Manila, December (1989) WMO/TD-No. 319 WMO, Geneva. 277-293

17 T S Murty *Canad Bull Fisheries and Aquatic Sciences, Bull* (1984) 212

18 Sündermann J and W Lenz *North Sea Dynamics* Springer-Verlag Berlin Heidelberg (1983) 693pp

19 WMO (1990) WMO Second International Workshop on Tropical Cyclones (IWTC-110, WMO/TD No. 83, WMO, Geneva)

20 A Sielecki *Mon Wea Rev* **96** (1968) 150-156

21 B Johns, P C Sinha, S K Dube, U C Mohanty and A D Rao *Q J R met Soc* **109** (1983) 211-224

22 N S Heaps *Geophys J Astron Soc* **35** (1973) 99-120

THE NOR'WESTER SCENARIO: THE ROLE OF INTERNAL GRAVITY WAVE

D K SINHA AND G K SEN
Centre for Atmospheric Sciences, University of Calcutta,
92, *Acharya Prafulla Chandra Road, Calcutta* 700 019, *India*

(*Received* 15 *June* 1993; *Accepted* 17 *June* 1993)

The understanding of the development of Nor'westers, through the mechanism of momentum and energy transfer of Internal Gravity Wave, forms the chief interest of this paper. A simulation experiment has been done to show that due to such transfer processes with changing phase velocity of disturbances, lower stable layer turns out to be unstable with significant growth as phase velocity approaches the true storm speed of that day. The findings also reveal the role of upper and lower layers in storm-triggering processes.

Key Words: Nor'Westers Scenario; Internal Gravity Wave; Momentum; Energy Transfer; Storm-triggering

Introduction

Nor'westers are known to be severe storms that are dominantly characteristic of the eastern part of India occurring usually between March and preonset phase of the Indian summer monsoon. These storms usually come from Northwest, being developed deeply in Chotanagpur belt, *vide*, *Tech. Rep.* IMD No. 10 (1941)[1]. This storm travels along Northwest to Southeast direction with a speed of 30 to 40mph and is normally accompanied by strong surface winds, hailstorms and often tornadoes. The extant literature on such storms says that a very tall cloud growth is an associated feature; it is a mesoscale process where horizontal scale is comparable to vertical scale; in fact, according to Rao and Mukherjee[2], there can be cloud development upto an altitude of 20km. This storm can last from 30 min. to several hours; Koteswaram and Chakraborty[3] pointed out the importance of Chotanagpur belt of Bihar plateau in initiating vigorous convection, which is apparent from the development of convective clouds over the slopes of high terrain, Ramaswamy[4], through his extensive analysis of sea level and upper air chart, pointed out that the dynamics of middle and upper tropospheric west wind belt and thermal advection process are more potent to Nor'westers. He also found wave pattern in jet stream which is a conspicuous feature of 700-500mb and 500-300mb partial thickness line. Koteswaram and Srinivasan[5] mentioned that development of Nor'westers associated thunderstorm is possible due to superposition of upper divergence associated with a straight jet, anticyclonic vortex on convergence associated with low level flow, even southernly flow. As most of the thunderstorms start in the afternoon, the role of insolation and orography in triggering thunderstorms is also an additional feature. These investigations lead us to the fact that Nor'westers, having originated in areas close to Chotanagpur belt of Bihar, chiefly occur on account of the strong influence of convection; and its dynamics is controlled by strong wind shear at lower and upper troposphere, with a strong moisture content at lower levels. The time of occurrence of such storms, their speed and its horizontal length scale, continue to be basic concerns for in-depth studies of Nor'wester phenomena, particularly the understanding of physical processes involved.

In atmospheric phenomena concerning storms, it is customary to seek understanding of the physics of the development of storms from different standpoints. One kind of investigation in respect of storm generation is to study, how an Internal Gravity Wave (IGW) influences the transport of energy and momentum from higher to lower levels leading to an augmentation of wind speed, thereby increasing instability at lower levels, *vide* Gossard and Hook[6], King *et al.*[7], Lalas and Einaudi[8], Lin and Goff[9], Lindzen[10], Gill[11] and Gossard and Moninger[12]. To the best of knowledge of the authors, the study of development of Nor'westers has not been looked into so far from this standpoint. The purpose of the present paper is to establish, from theoretical standpoint reinforced by available data, the existence of Internal Gravity Wave (IGW) thereby facilitating understanding of possible mechanisms for transport of momentum and energy to the lower levels. The layout of this paper is as follows.

To begin with, a description of the model of internal gravity wave drawing upon the works of Lindzen[10] and Gill[11], is set forth. Then comes the method of solution followed by an analysis, from theoretical considerations, of the estimation of parameters. Next to provide a realistic flavour of this investigation on propagation of internal gravity wave affecting stability or otherwise in subsequent stages, we perform a simulation experiment on the basis of IMD data.

A critical analysis of profiles obtained from the experiment provides us a better insight about the processes of energy and momentum transfers through IGW leading to conspicuous instability at lower levels.

Model Description

For a basic flow with vertical shear $\left(\frac{\partial u}{\partial z}, \frac{\partial v}{\partial z}\right)$ in the z-direction, the equations, of motion, continuity and energy conservation are, *vide*, Lindzen[10], Gill[11].

$$\rho\left[\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right]u=-\frac{\partial p}{\partial x}, \qquad \ldots (1)$$

$$\rho\left[\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right]v=-\frac{\partial p}{\partial y}, \qquad \ldots (2)$$

$$\rho\left[\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right]w=-\frac{\partial p}{\partial z}-g\rho \qquad \ldots (3)$$

and

$$\left.\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right\}\rho+\rho\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right. = 0\,. \qquad \ldots (4)$$

Taking the atmosphere to be a Boussinesq fluid, we can split eq. (4) into two equations:

$$\left\{\frac{\partial}{\partial t}+u\frac{\partial}{\partial x}+v\frac{\partial}{\partial y}+w\frac{\partial}{\partial z}\right\}\rho=0\,. \qquad \ldots (5)$$

(under the incompressibility condition)

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \qquad \ldots (6)$$

Here u, v, w, ρ, p are zonal velocity, meridional velocity, vertical velocity, density and pressure respectively. The independent variables are time (t), zonal direction (x), meridional direction (y) and vertical direction z. As usual, we now expand the state variables into basic and perturbed parts.

$$\begin{aligned}
&u=u_0(z)+u'(x,y,z,t)\,,\\
&v=v_0(z)+v'(x,y,z,t)\,,\\
&w=w'(x,y,z,t)\,, \qquad \ldots (7)\\
&\rho=\rho_0(z)+\rho'(x,y,z,t)
\end{aligned}$$

and

$$p=p_0(z)+p'(x,y,z,t)\,.$$

It is to be remarked, in passing, that both meridional and zonal velocities are important for Nor'wester development. Here, we have retained both the unperturbed velocities for u and v, disturbances being from North-Western side to travel along South-Eastern direction. We next construct the linearized equation and then adopt the following transformation of perturbed variables as

$$u',\ v'\ \ w',\frac{1}{p'}\rfloor\ \ \left(\frac{\rho_s}{\rho_0}\right)^{1/2}\ \ U,\ V,\ W,\frac{1}{P} \qquad \ldots (8)$$

The unperturbed vertical velocity, being negligible, is omitted. This is also warranted by our subsequent simulation experiment which presupposes a stable state of thermal stratification.

Here $\rho_0(z)$ is the unperturbed density and ρ_s is the density at surface.

The resulting partial differential equation, is linear in U, V, W and P'; we seek its solution being in the form:

$$\begin{aligned}
&U=U(z)\sin(kx+ly-\sigma t);\\
&V=V(z)\sin(kx+ly-\sigma t); \qquad \ldots (9)\\
&W=W(z)\cos(kx+ly-\sigma t);\ \text{and}\\
&P=P(z)\sin(kx+ly-\sigma t).
\end{aligned}$$

Elimination of $U(z)$, $V(z)$ and $P(z)$ leads to the ordinary differential equation in $W(z)$:

$$\frac{\mathrm{d}^2 W}{\mathrm{d}z^2} + \alpha(z)\, W(z) = 0\,, \qquad \ldots (10)$$

where

$$\alpha(z) = -m^2 + \frac{1}{\omega}\left(k\frac{\mathrm{d}^2 u_0}{\mathrm{d}z^2} + l\frac{\mathrm{d}^2 v_0}{\mathrm{d}z^2}\right) + \left(\frac{Nm}{\omega}\right)^2$$

$$\omega = k\left(C - u_0 - \frac{l}{k} v_0\right)$$

and

$$m^2 = l^2 + k^2 \quad C = \frac{\sigma'}{k}\,, \; N^2 = -\frac{g}{\rho_s}\frac{\mathrm{d}\rho}{\mathrm{d}z} = \frac{g}{\theta_v}\frac{\mathrm{d}\theta_v}{\mathrm{d}z} \qquad \ldots (11)$$

Here l and k are horizontal wave numbers along zonal and meridional direction, σ is the wave angular frequency. C is the phase speed. N^2 is the Brunt-Vaisala frequency, θ_v is the virtual potential temperature, ω is the expression for intrinsic frequency.

The solution of eq. (10) is sought for a multilayered model. At the surface, assumed to be a rigid level, we set $W=0$. At the interface between several layers, the kinematic condition is $W_L = W_U$ and dynamic condition is $P_L = P_U$, *where* L represents lower and U upper layer, *vide*, Gossard and Hook[6]. It is to be noted that both w and $\mathrm{d}w/\mathrm{d}z$ are continuous at the interface. This is also true for the undisturbed interface.

Method of Solution

We first compute $\alpha(z)$ at several grid points along vertical, where measured values of atmospheric variables are available from Radiosonde data, *vide* IMD[1]. After computing the values of $\alpha(z)$ we divide the atmosphere into several layers where in each layer $\alpha(z)$ profile is either concave or convex. From each layer, we choose three grid points one at the bottom of layer, one at the top of layer and another, in between where $\alpha(z)$ attains either minimum or maximum value i.e., $\frac{\mathrm{d}\alpha}{\mathrm{d}z} = 0$. Using the three values of $\alpha(z)$ at these three points, we try to fit them using a function of the form

$$\alpha(z) = -\frac{\mu^2}{4} + \frac{E^2}{(1 + D_1 e^{\mu z} + D_2 e^{-\mu z})^2} \qquad \ldots (12)$$

So in each layer, we have an estimate of μ, E, D_1 & D_2 respectively. Then we also estimate K_r and K_i to be used later on, from the relation

$$K^2 = E^2 + \mu^2(D_1 D_2 - 1), \qquad \ldots (13)$$

where $K_r=\sqrt{K^2}$ if $K^2>0$ and $K_i=\sqrt{-K^2}$ if $K^2<0$.

For $K^2>0$ i.e., $K_r\neq 0$ and $K_i=0$ the final solution is of the form

$$W=\sqrt{|1+D_1e^{\mu z}+D_2e^{-\mu z}|}\,(C_1\sin\theta+C_2\cos\theta), \qquad \ldots (14)$$

where

$$\theta=\frac{-2(K_r+K_i}{\mu\sqrt{4D_1D_2-1}}\left\{\tan^{-1}\left(\frac{1+2D_1e^{\mu z}}{\sqrt{4D_1D_2-1}}\right)-\tan^{-1}\left(\frac{1+2D_1e^{\mu z_0}}{\sqrt{4D_1D_2-1}}\right)\right\}, \qquad \ldots (15)$$

when

$4D_1D_2-1>0$

and

$$\theta=\frac{(K_r+K_i)}{\mu\sqrt{4D_1D_2-1}}\left\{\log_e\frac{(1+2D_1e^{\mu z})-\sqrt{1-4D_1D_2}}{(1+2D_1e^{\mu z})+\sqrt{1-4D_1D_2}}\right.$$

$$\left.-\log_e\frac{(1+2D_1e^{\mu z_0})-\sqrt{1-4D_1D_2}}{(1+2D_1e^{\mu z_0})+\sqrt{1-4D_1D_2}}\right\} \qquad \ldots (16)$$

when

$4D_1D_2-1<0$

for $K^2<0$ i.e., $K_r=0$ and $K_i=0$,

$$W(z)=\sqrt{|1+D_1e^{\mu z}+D_2e^{-\mu z}|}\,[C_1\sinh\theta+C_2\cosh\theta]. \qquad \ldots (17)$$

The expressions of θ, are same, as described by (16) and (15) Z_0 is the height of lower boundary of each layer.

Our next task is to match the values of $w(z)$ and $\frac{dw}{dz}$ at the interface of each layer. It is also important to satisfy the lower boundary condition. This enables us to estimate the values of constants c_1 and c_2 in each layer.

Simulation Experiment: Analysis and Discussion

As stated earlier, this experiment is based on IMD data set. The period of observation is from May 10, 1988 to May 25, 1988. The data consisting of usual radiosonde observations, are taken 4 times daily at 5-30 IST, 10-45 IST, 17-30 IST and 22-45 IST at five stations Calcutta, Bhubaneswar, Ranchi, Patna and Siliguri. On 22-05-1988 at Calcutta a Nor'wester storm occurred at 19-40 IST. The characteristic horizontal length scale of this disturbance is 20km. The data we have used here is the radiosonde observation taken at 0 GMT or 5-30 IST at Calcutta on 22-05-1988, which is 14 hours prior to storm. Let us now have an indepth analysis on the stability or otherwise of the internal gravity wave on the basis of modelling, done earlier and simulation. This allows the phase ve-

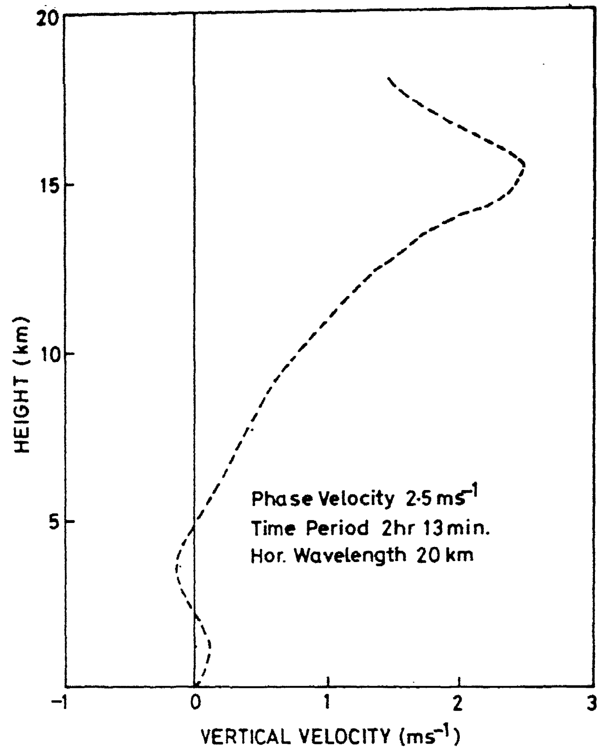


Fig 1

locity to assume appropriate values sequentially close to the known value of the storm speed, as available from IMD data. What we seek precisely here is to get at positive or negative values of K^2, given by eq. (13), enabling us to have a qualitative information about stability or instability of the layer concerned. In mathematical terms, this is an eigenvalue problem. Accordingly, the simulation experiment is structured as follows:

In order to analyse our findings, we draw the profiles of perturbed vertical velocity versus height. Fig. 1, provides the graph of the first experiment with phase velocity 2.5 ms^{-1}; one also finds internal gravity waves with wave length 4.5km, obviously well below 5km indicative of its existence in lower tropospheric levels. The 5-15km level, shows a growth of amplitude with its maximum at 15km. It is found that K^2 is negative throughout this layer. Hence, any perturbation at 15km. level will have the maximum amplitude and propagate downward as internal gravity wave. This is generated because of a horizontally propagating disturbance with length scale 20km and speed 2.5ms^{-1}.

The second step in the experiment is to increase the phase velocity to 5.55ms^{-1} and the profiles are drawn in Fig. 2. The internal gravity wave in this case is of a very small amplitude and vertical wavelength 4km. A growth on account of instability is observed in the layer 4-13km. Then follows a remarkable growth (through negative amplitude increment) with a maximum at 16.5km. In passing, one finds strikingly a negative amplitude, contrary to posi-

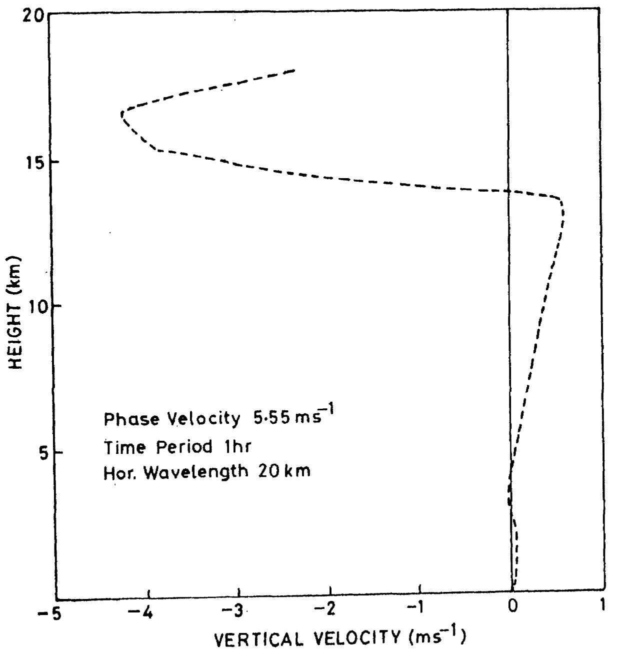

HEIGHT (km)
20
15
10
5
VERTICAL VELOCITY (ms^{-1})
-5
-4
-3
-2
-1
0
1
Phase Velocity 5·55 ms^{-1}
Time Period 1hr
Hor. Wavelength 20 km

Fig 2

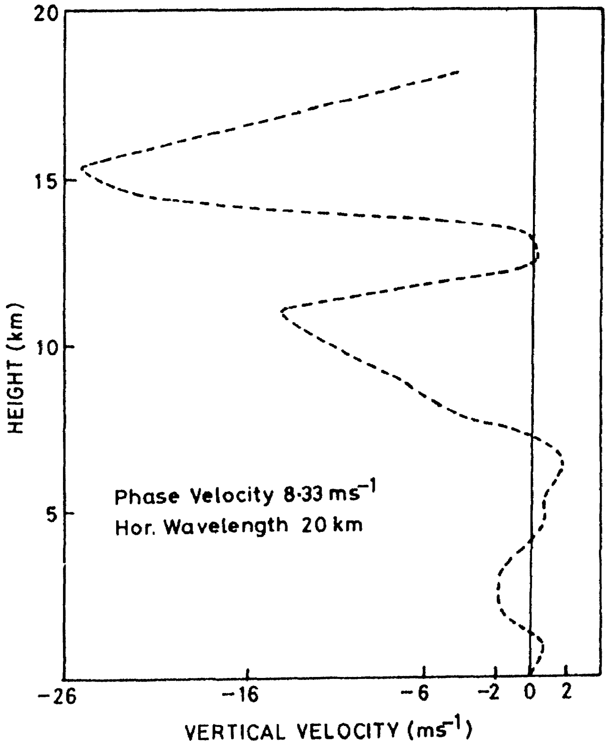

HEIGHT (km)
20
15
10
5
VERTICAL VELOCITY (ms^{-1})
-26
-16
-6
-2
0
2
Phase Velocity 8·33 ms^{-1}
Hor. Wavelength 20 km

Fig 3

tive amplitude in the earlier experiment above 13km level. Further, the updraft zone has decreased in height with the increase of phase speed, and one can thus call it a strong downdraft zone, roughly above 14km, which was not observed in the first experiment.

The phase velocity of the third experiment (Fig. 3) is 8.33ms^{-1} or 30km. per hour. Here $K^2 < 0$, for $Z < 2$km, $7 < Z < 12$km, $13 < Z < 15$km. The internal gravity wave generated at 15km. level, as before, propagates downward. The modes at several other layers being unstable, it is likely that the internal gravity wave generated at 11km level may also propagate downward. It thus looks possible that the internal gravity wave generated in the level $13 < Z < 15$km in its downward propagation, has its amplitude augmented in the unstable layer between 7 and 12km, strictly at 11km level, even though it may suffer otherwise. The vertical wavelength is 4km. The downdraft zone is clearly within $7 < Z < 12$km. One may think in terms of possibility of cloud in this layer of 5km thick.

In the final experiment, the phase velocity is 8.9ms^{-1}·32km hr^{-1}. The outcome is shown in Fig. 4. A strong downward draft from 1km to 2.3km is observed. Further, below, 1km. a weak updraft is observed, which characterises a feature of weak convection before storm, as as is obvious from 1730 sounding.

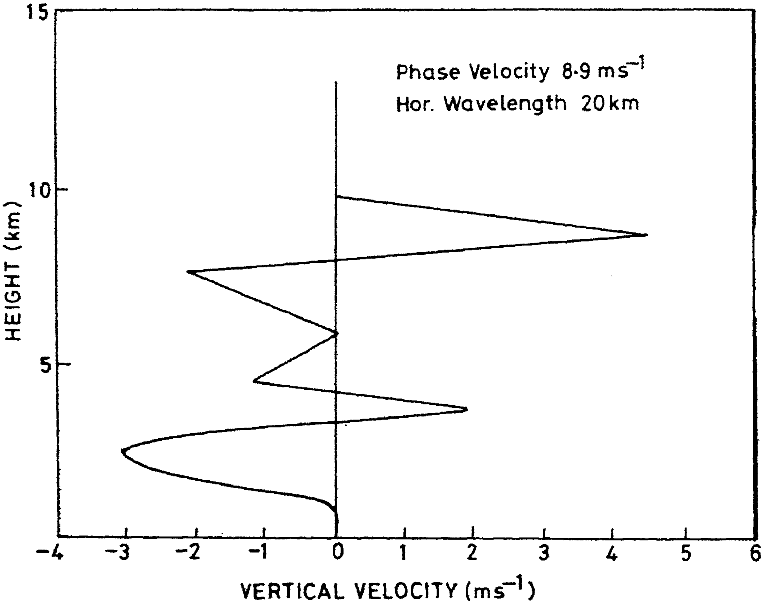


Fig 4

A comparative study of all the four experiments may now be undertaken. First, from Figs 1 and 2, we find that, for a low phase speed of propagation, the Internal Gravity Wave with vertical wavelength 4 to 4.5km is observed. In such a case, phase speed is below 5.5ms^{-1}. Second, in all the cases, Internal Gravity Wave is generated at 15km level, due to the existence of instability. Third, while the increase in phase speed makes all gravity wave modes more and more unstable, the increase in amplitude takes place at lower levels. Specifically speaking, with increased value of the phase speed and constant horizontal wavelength, the lower troposphere is found to exhibit more instability than what it was at low phase speed. Hence, it is surmised that at higher phase speed of disturbance, the momentum and energy transport by Internal Gravity Wave, which is, of course, generated at higher unstable layer, are increased. Finally the lower troposphere becoming unstable as disturbance attains phase speed of 32km hr^{-1}, one observes a beginning of strong downdraft which becomes often obvious before storms.

Conclusions

The foregoing investigation, even though essentially theoretical in nature, brings out an interesting feature of generating IGW between 4-13km layer because of instability. While this is not somewhat unusual in respect of IGW in the context of severe thunderstorm as observed through satellite observation on cloud elsewhere *vide* Ley and Peltier[13], findings here in respect of Nor'westers show a striking resemblance to the observational findings on the existence of simply waves in that form as observed by Ramaswamy[4]. Further, that instability can occur in a few specific layers, namely 4-11km layers is also revealed from this investigation, and this finds a fairly close agreement with what have been indicated by Koteswaram and Srinivasan[5].

Acknowledgement

Authors wish to thank the Department of Science & Technology, Government of India for financial support and IMD for data provided to them.

References

1 India Meteorological Department (IMD) (1941) *Nor'westers of Bengal Tech Note* No. 10
2 D V Rao and A K Mukherjee *Indian J Met Geophys* **I** (1958) 313-322
3 P Koteswaram and K C Chakraborty *Indian J Met Geophys* I (1950) P-167
4 C Ramaswamy *Tellus* **8** (1956)
5 P Koteswaram and V Srinivasan *Indian J Met Geophys* **9** (1958) 301.
6 E E Gossard and W H Hook *Waves in the Atmosphere* Elsevier Scientific Publication, New York (1975) 456
7 J C King, S D Mobbs, M S Derby and J M Rees *Bound Layer Met* **39** (1987) 1-13
8 D P Lalas and F Eissandi *J appl Met* **13** (1974) 318-324
9 Y H Lin and R C Goff *J atmos Sci* **45** (1988) 194-205
10 R S Lindzen *Dynamics in Atmospheric Physics* Cambridge University Press, Cambridge (1990)
11 F A Gill *Atmosphere Ocean Dynamics* Academic Press, New York (1982)
12 E E Gossard and W R Moninger *J atmos Sci* **32** 2111-2123
13 E B Ley and W R Peltier *J atmos Sci* **38** (1981) p1206

LONG RANGE WEATHER FORECASTING TECHNIQUES IN INDIA USING EOF APPROACH

H N SRIVASTAVA
India Meteorological Department, Pune 411 005, *India*
AND
S S SINGH
India Meteorological Department, New Delhi 110 003, *India*

(*Received* 21 *June* 1993; *Revised* 3 *September* 1993;
Accepted 6 *September* 1993)

Empirical Orthogonal Functions (EOF) have been used to study the interdependence of 16 parameters in power regression and parametric models used by the India Meteorological Department for predicting the summer monsoon rainfall. Alternate models, using an eigen index have been discussed for the country as a whole based on 16 parameters, and for the onset of the monsoon over Kerala using ten other parameters. Distinctive aspects during excess and drought years (± 10%) have been brought out. An attempt has also been made to evolve models for smaller areas using rainfall correlations with eigen indices enabling us to demarcate three broad regions in the country.

Key Words: Long Range Forecast; Empirical Orthogonal Functions; Principal Component Analysis; Southwest Monsoon

Introduction

Seasonal monsoon (June to September) rainfall has a direct bearing on the Indian economy. Attempts have, therefore, been made by the India Meteorological Department to predict monsoon rainfall since 1886, when the first operational Long Range Forecast (LRF) was issued by Blanford[1]. Subsequently, multiple regression techniques based on regional and global data have been used starting with the work of Walker[2], who evolved the concept of a southern oscillation—a seesaw fluctuation of atmospheric pressure from the Indian to the Pacific ocean. The parameters in multiple regression models were changed due to changing correlations between one or the other with the progress of time. Keeping in view these limitations, Thapliyal[3] introduced a dynamic stochastic model which was also used for operational LRF for peninsular India. However, limitations were noticed as the method was based on only one parameter, namely, the position of a 500hPa ridge over peninsular India first reported by Banerjee *et al.*[4], Gowarikar *et al.*[5], evolved a 'parametric model', wherein the favourable parameters out of a total of 16 are expressed as a percentage to, qualitatively, forecast monsoon rains over the country. Gowarikar *et al.*[6] also developed a power regression model based on the same global and regional parameters to provide a quantitative approach. The authors felt that some of the parameters are inter-related, but without presenting a scientific analysis. Keeping in view the considerable discussion among scientists about this aspect, Srivastava and Singh[7] employed Empirical Orthogonal Functions (EOF) to evaluate the interdependence of the parameters in these two models.

This technique was earlier used by Bedi and Bindra[8] to explain different rainfall patterns over India. Iyengar[9] also applied the method to understand the variability of rainfall in Karnataka.

The aim of this paper is to discuss the interdependence of the LRF parameters, and to present models using EOF to predict seasonal rainfall for the country as a whole and for smaller regions using additional data. A model for the monsoon's onset over Kerala has been also proposed using the same technique.

Methodology

Principal Component Analysis (PCA) has been extensively used in Meteorology after Lorenz[10] introduced the technique in 1956. It expresses a meteorological parameter, such as rainfall as a linear combination of orthogonal functions. Expressing different meteorological parameters *m* for a series of *n* years, the eigenvectors of the covariance matrix are computed. The Empirical Ortho-

Table I

Parameters used for long range forecast of southwest monsoon

(*a*) Onset

(1) Cobar (Australia), November zonal wind (500hPa)
(2) Thiruvananthapuram, Madras, February wind (200hPa)
(3) Darwin (Australia), January wind (200hPa)
(4) Delhi, January wind (300hPa)
(5) Forrest (Australia), January wind (200hPa)
(6) Thiruvananthapuram, November kinetic energy (200hPa)
(7) Fukuoka (Japan), February kinetic energy
(8) Mean January zonal wind over Darwin (Australia) (300hPa)
(9) Mean meridional December wind (previous year) over Calcutta (200hPa)
(10) Kinetic energy of January over Darwin Australia) (200hPa)

(*b*) Rainfall

(1) 10hPa westerly winds over Balboa (January)
(2) Darwin pressure (April)
(3) Eurasian snow cover (previous December)
(4) El-Nino in current year (October to previous year to May of current year)
(5) Equatorial pressure (January to May)
(6) Himalayan snow cover (January to March)
(7) Location of 500hPa ridge over India (April)
(8) Central India minimum temperature (May)
(9) 50hPa ridge-trough extension over northern hemisphere (January and February)
(10) North India minimum temperature (March)
(11) East coast minimum temperature (March)
(12) Northern Hemispheric temperature (January and February)
(13) El-Nino in previous year
(14) Argentina pressure (April)
(15) Tahiti-Darwin pressure (March to May)
(16) Northern hemispheric pressure (January to April)

gonal Function (EOF) is the set of coefficients appearing in the first principal component which is the linear function explaining maximum variance. To define the principal components uniquely, the eigenvalues are normalised (Preisendorfer[11]).

Long Range Forecasting Parameters Over India

Das[12] reported Long Range parameters for the onset date as well as the quantum of rainfall in multiple regression models used by the India Meteorological Department. These are constantly reviewed and changed with the passage of time. Table I gives a list of recent predictors for predicting the monsoon onset date and the southwest monsoon rainfall.

Results

Table II provides an analysis of first 4 EOFs, which explain more than 62% of rainfall over the country. Bedi and Bindra[8] however, reported only 47% of variance of 'monsoon rain' through the first 4 EOF's. However, the first EOF which explains 25.9% variance is in agreement with the earlier study. A detailed analysis was undertaken (Table III) for the excessive and deficit monsoon rainfall ($>$ or $<10\%$ of normal rainfall) to find out any distinctive aspects. Table IV presents an analysis for the onset date over Kerala. The correlation matrix of subdivisional monsoon rainfall and EOF for three regions, namely, SE India, NE India and India-(NE + SE) was also worked out and given in Table V.

Eigen Indices corresponding to the meteorological parameters for the onset date as well as southwest rainfall were computed separately from the linear combination of their respective values and the loading factors for each year are interpreted with respect to the actual onset date or the quantum of rainfall (Srivastava and Singh[7]).

Discussion

We note from Table II that the first EOF shows a high correlation of 0.83 with the summer monsoon rainfall for the country if we consider the data of 34 years (1958-91). It accounts for a variance of about 26%. The parameters which make a singificant contribution to explain the variability were found to be (*i*) 50hPa E-W ridge, (*ii*) Eurasian snow cover, (*iii*) 500hPa ridge, (*iv*) 10hPa zonal wind, (*v*) Central India temperature, (*vi*) east coast temperature, (*vii*) northern India temperature, (*viii*) 40° northern hemisphere pressure, and (*ix*) Argentina pressure. Four other parameters, namely, Northern Hemisphere temperature, Equatorial pressure, El-Nino of the current year and the Darwin pressure had relatively lesser influence. The southern oscillation index (Tahiti minus Darwin pressure) and the Himalayan snow cover were found to have least loading followed by the El-Nino of previous year.

Paranjpe and Gore[13] have used a parsimonious logistic regression model and found that only 5 variables namely the 500hPa ridge, 10hPa wind with temperature in northern hemisphere, the temperature in north India, the tem-

Table II

Coeffi *ors for first four EOFs*

Serial Number	E-WRIG	EURASIONSNOW	RIG. 500 MB	ZONAL 10 MB	VID IN MAY	CVMM TN MAR	PRESS 40 J	N-H TEMP J	ARGET PRS	TD PRS	ELNI(P)F	EUQAT PRS	NORT.I.TEM	ELNI(C)	DAR. PRS	HIMAL. SNOW	VARIANCE EXP. %	CORRELATION
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	0.38	0.38	0.31	0.31	0.26	0.29	0.30	0.18	0.31	0.07	0.10	0.16	0.25	0.14	0.15	0.08	25.98	0.83
2.	0.00	0.02	0.08	0.01	0.11	0.25	0.04	0.31	0.10	0.45	0.17	0.47	0.17	0.17	0.46	0.31	18.69	0.01
3.	0.42	0.11	0.19	0.02	0.18.	0.12	0.19	0.29	0.09	0.25	0.30	0.01	0.43	0.25	0.06	0.44	9.14	0.27
4.	0.00	0.01	0.02	0.09	0.29	0.14	0.26	0.11	0.25	0.20	0.56	0.06	0.16	0.60	0.07	0.08	7.15	−0.15

Table III

EOF Analysis of excess and deficit monsoon years (± .10%)

(*a*) *Deficit Monsoon Years* (1965, 1966, 1972, 1974, 1979, 1982, 1986 &1987)

SERIAL NUMBER	ARGET. PRS.	CINDIA TEMP.	NORT.I.TEMP.	ZONAL 10 MB	T-D SPR.	E.W.RIG.	ELNI(P)	EURASION SNOW	EQUAT. PRES.	ELNI(C)	NH PRESS 40°N	RIG 500 MB	HIMAL. SNOW	C.I. TEMP. MAY	DAR. PRES.	NH TEMP.	VAR. EXPLAINED	CORRELATION
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	0.12	0.17	0.01	0.01	−0.26	−0.08	0.22	0.20	0.32	0.37	0.16	−0.40	−0.28	−0.23	0.27	0.40	33	.02
2.	−0.47	−0.36	0.01	−0.06	0.34	0.01	0.04	−0.18	0.31	0.22	0.30	0.12	0.25	−0.34	0.26	−0.07	23	.42
3.	−0.17	−0.19	−0.30	−0.48	0.10	0.47	0.21	0.28	0.13	0.07	−0.38	0.04	−0.24	0.19	0.03	−0.02	18	−.84
(*b*) *Excess Monsoon Years* (1959, 1961, 1970, 1975, 1983 &1988)																		
1.	0.11	−0.22	0.04	−0.25	0.24	0.75	−0.05	−0.11	−0.30	0.17	0.16	−0.28	0.39	0.33	0.35	0.37	30.11	−.47
2.	−0.24	0.10	0.35	0.17	−0.34	0.30	−0.34	−0.30	0.25	0.27	0.34	0.13	0.08	0.26	0.05	0.21	25.77	−.43
3.	0.42	−0.47	−0.04	−0.21	−0.06	0.21	−0.23	0.29	0.29	0.35	−0.18	0.02	−0.13	−0.06	0.31	−0.06	16.51	−.50

Table IV

Eigen vector for long range prediction of summer monsoon onset

Eigen vectors of parameters												
EOF Parameters	COBAR WIND	THIRU.-MADRAS WIND	DARWIN WIND	DELHI JAN. WIND	FORREST (AUS) JAN. WIND	THIRU.NOV. KINETIC ENRG	FUKUOKA (JAPAN) FEB, KIN ENRG.	MEAN JAN.WIND DARWIN	MEAN MERID. DEC. WIND	KINE. ENRG. JAN. DARWIN	VARIANCE EXPLAINED	CORR. COEF. WITH THE ONSET DATES SW MONSOON OVER KERALA COAST
1	2	3	4	5	6	7	8	9	10	11	12	13
1.	0.46	0.17	−0.43	−0.17	0.33	0.20	0.13	−0.37	0.35	0.35	31.13	0.79
2.	0.15	0.45	0.33	−0.53	0.34	−0.16	−0.06	−0.15	−0.26	−0.38	20.45	0.31
3.	0.15	−0.29	0.28	0.00	−0.04	−0.12	0.70	−0.09	0.40	−0.37	13.75	0.41
4.	0.04	−0.13	−0.09	−0.27	0.48	−0.79	0.00	0.08	−0.05	0.18	11.70	−0.21

Table V

EOF correlation matrix and broad sub-divisions in India

Broad sub-divisions	Correlation matrix of broad sub-divisional monsoon rainfall and EOF															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
India-(NE&SE)	.84	.02	.14	.18	.18	.01	.01	.15	.14	.21	0.40	.09	.05	.00	.13	0.7
SE India	.36	.18	.28	.39	.03	.04	.16	0.60	.01	.43	.10	.14	.20	.06	.19	.20
NE India	.04	.31	.18	.02	.27	.08	.13	0.70	.23	.31	0.06	.37	.17	.18	.00	.04

perature of the east coast and the interaction between temperature in north India and 10hPa wind may be used to assess the probability of a normal rainfall over the country. Keeping in view the advantage of EOF techniques, whereby the complicated variability of the original data set is reduced to a relatively few uncorrelated components, we found that an eigen index gives similar results as the parametric model on 25 occasions to bring out the correct rainfall. For three years, namely, 1967, 1978 and 1985, the actual rainfall could not be predicted by either of the two methods. During the years 1963 and 1966, a negative index supported deficient rainfall. This suggested an improvement over the parametric model.

It is interesting to note that the second EOF which explains about 19% of the variance has significant influence on the southern oscillation, equatorial pressure and the Darwin pressure, followed by the northern hemispherical temperature, Himalayan snow cover and the east coast temperature. However, the correlation with monsoon rainfall was found to be negligible. The influence of the third and the fourth EOF was about 9% each which gave correlations of 0.27 and −0.15 respectively, thus explaining the total variance of 62% through the first four EOF's. It may however, be noted that Shukla and Paolino[14] and Prasad and Singh[15] found the Darwin pressure tendency from winter to spring to be a good indicator of monsoon rainfall. Although IMD is using Darwin pressure (spring) instead of its pressure tendency, the use of one or two parameters to predict seasonal rainfall has not been successful in the past. Parthasarthy *et al.*[16], using 11 circulation parameters (5 parameters used by Gowariker *et al.*[6] and 6 new parameters) found that three EOF's explain 69% variance with a multiple correlation of 0.82.

By considering the excess and deficit monsoon years (rainfall more than or less than 10% of the normal) separately, it may be noted from Table III that the 3rd EOF shows a significant correlation of −0.84 during the deficit monsoon years. The variance explained was only 18%. But if the second EOF is included, which shows a correlation of −0.42, the variance explained increased to 41%. On the other hand, during excess monsoon years, first three EOF's explain about 72% of the variance and their correlations with rainfall were −0.47, −0.43 and −0.50 respectively. Considering the first EOF during excess monsoon years and in common with drought years based on the third EOF (Table III), it may be noted that northern India temperature, Eurasian snow cover, northern Hemispheric pressure and the El-Nino of the previous year influences the deficit rainfall years. On the other hand, central India and east coast temperatures, equatorial pressure, 500hPa ridge, northern hemispheric temperature, Tahiti-Darwin pressure and Darwin pressure influence excess monsoon years. However, the poor correlation with the first EOF during deficit monsoon years and its exclusion from the above discussion needs to be kept in view.

An attempt was made to identify smaller areas in the country after studying the correlation coefficients of monsoon rainfall with the eigenvectors in different meteorological sub-divisions (Singh *et al.*[17]). The meteorological sub-divisions were grouped together which showed a significant correlation (at 5% level of significance) with the EOF's. It may be noted that although EOF's were

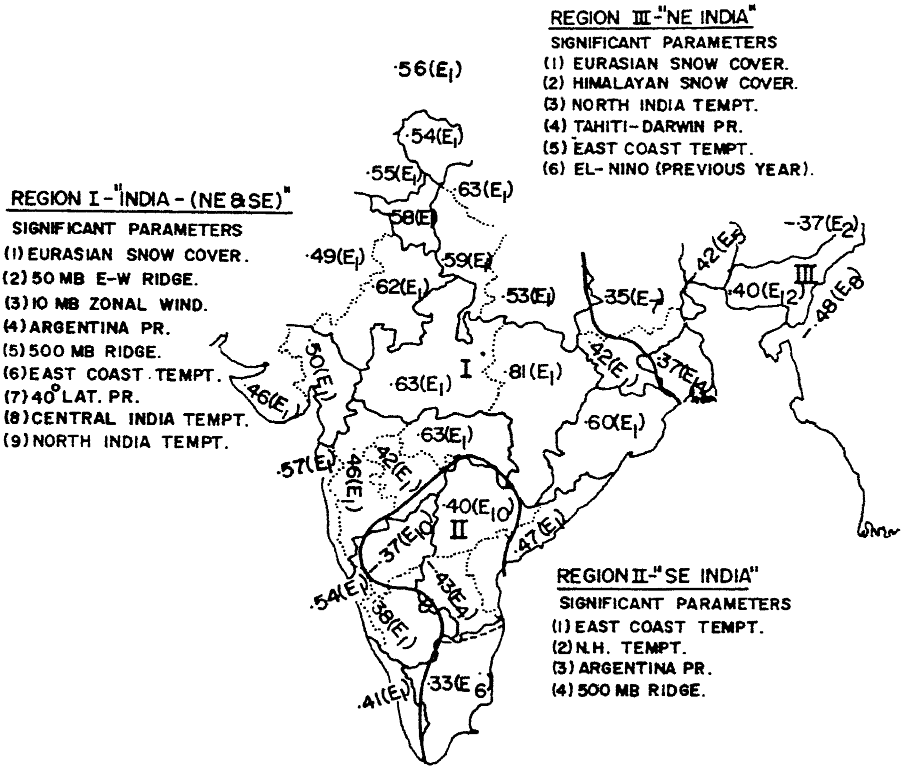


Fig 1 Broad divisions in India based on EOF analysis for Long Range Forecasting

computed for all the parameters, some of the parameters which do not make a significant contribution in the variability of the function were excluded. Thereafter, different regions of the country and the corresponding EOF's were also identified (Fig. 1) such that the interannual variability patterns of the monsoon rainfall and the EOF's were found similar to each other over the respective region. The results are shown in Fig. 1. After detailed analysis of the first EOF and their linear functions, it was found that the main parameters to explain interannual variability in total rainfall during the monsoon season over sub-division I namely, India minus (NE + SE) parts, are the 10hPa zonal wind, 500hPa ridge, 50hPa east west ridge, Vidarbha temperature, east coast temperature, North India temperature, Eurasian snow, Argentina pressure and pressure at 40°N latitudinal belt.

It was interesting to note that the first EOF was not significantly correlated with Tamil Nadu rainfall during this season which provides weightage to the methodology, keeping in view the climatological and synoptic pattern of rainfall. EOF 4 and EOF 10 for the broad sub-division II explained interannual variability through east-coast temperature, northern hemisphere temperature, Argentina pressure and the 500hPa ridge. In the northeastern part of India, EOF 7 (Bihar Plain), EOF 14 (Gangetic West Bengal), EOF 2 (Arunachal Pradesh), EOF 12 (Assam and Meghalaya), EOF 5 (Sub Himalayan Gangetic

West Bengal), and EOF 8 (Nagaland, Manipur and Tripura) gave the following predominant factors: as (*i*) Tahiti-Darwin pressure, (*ii*) North India temperature, (*iii*) East coast temperature, (*iv*) Eurasian snow, (*v*) Himalayan snow and (*vi*) El-Nino (previous year).

It may be mentioned that Bedi and Bindra[8] have shown an opposition in the rainfall pattern between northeast and western sectors of India. The present study brings out that maximum correlation has been found with the first EOF in northwest India while over the northeast India EOF's 2, 5, 7, 8, 12 and 14 show higher correlation. Expressed through regional and global parameters, and excluding the common parameters, it may be noticed from Fig. 1 that the northwestern sector of India has predominant loading due to (*i*) 50hPa E-W ridge, (*ii*) 10hPa zonal wind, (*iii*) Argentina pressure, (*iv*) 50hPa ridge, (*v*) 40°N latitude Pressure, and (*vi*) Central India temperature. On the other hand, northeast India has loading influence due to (*i*) Himalayan snow cover, and (*ii*) Tahiti-Darwin pressure and El-Nino (previous year). It is interesting to note that Gadgil *et al.*[18] found that the variation of the monsoon rainfall over the Indian region is characterised by two spatial scales corresponding to about 11 and 31 zones based on EOF analysis. However, the authors suggested further studies to understand the mechanism for two scales in relation to the dynamics of the monsoon. As the present paper is based on the rainfall predictive parameters, the present study to divide the country into three broader zones appears to be justified.

EOF analysis was also applied for the onset of monsoon over Kerala based on parameters given in Table I. It shows that the eigen index derived from the first EOF gives 92% accuracy as compared to 80% based on the parametric method.

Future Developments

Although parameteric and power regression models have provided more accurate forecasts during the last 5 years but they are statistical in nature and in some year may provide divergent forecast (Gowariker *et al.*[5,6]). Considering the atmosphere as a complex dynamical system which transforms different inputs into outputs, such as the monsoon rainfall, transfer dynamic relationship between the continuous input and output series has been expressed by a differential equation[3], (given by seasonal rainfall) with 500hPa as the lead forcing to the atmosphere (Banerjee *et al.*[4]). During recent years, this model gave lesser accuracy, if we compare it with the power regression model. For example, the official forecast based on this model during 1990 and 1991 when compared with actual rainfall gave a difference of 10% and 7% during the last two years, which is contrary to the results of Thapliyal and Kulshrestha[19]. Considerable efforts to include other forcing parameters may be needed to improve the model. Another dynamical approach based on deterministic chaos found lower accuracy in prediction, which was attributed to limited rainfall observations (Kulkarni and Verma[20]).

Among the new parameters considered for the monsoon rainfall prediction over India could be sea surface temperature over the Arabian Sea, Bay of

Bengal and Indian Ocean. On extension of EOF analysis for these regions in relation to two contrasting sets of monsoon years, Balakrishnan *et al.*[21] found that higher eigenvectors in May over northeast Arabian sea may signal good monsoon and vice versa. Also the temperature gradient over the Bay of Bengal shows a contrast between excess and deficient monsoons for which satellite observations through OLR (outgoing longwave radiation) could provide data. A question arises whether a regional monsoon pattern could be influenced by persistent volcanic dust in the stratosphere emitted through volcanic eruptions similar to those of 1991 in the Phillipines. This aspect needs investigation in relation to LRF.

Development of long range forecast models for smaller areas over the Indian region, based on EOF analysis suggests that new predictive parameters for the country, excluding the northeast and southeast parts (Fig. 1) can be identified by the first EOF. However, in northeast India, where the rainfall variability is only 20% or slightly less, different EOF's have been found to correlate with the rainfall implying complexity in choosing predictive parameters. Considerable research is, therefore, needed to evolve models for smaller areas, namely, the three broad divisions in the country (Fig. 1) through the identification of more physically linked regional parameters.

The Climate Analysis Centre, Washington employs a modified statistical technique called Canonical Correlation Analysis (CCA) to predict Sea Surface Temperature (SST) for the central Pacific area, which has been found to be the most useful in long range prediction with 1 season, 2 seasons and 3 seasons lead forecasts (Ropelewski and Barnston[22]). However, their accuracy at present is slightly better or almost similar to numerical models. Improvements in coupled ocean atmosphere models and their testing over longer periods are still needed before they can be used for routine operational LRF.

Acknowledgement

The authors are grateful to Dr N Sen Roy, Director-General of Meteorology for permission to publish the paper. The authors also express their gratitude to Dr P K Das, FNA for helpful suggestions.

References

1 H F Blanford *Proc R Soc London* **37** (1884) 322
2 G T Walker *Mem India Met Dept* (1910) 21
3 V Thapliyal *Mausam* **33** (1982) 399
4 A K Banerjee, P N Sen and C R V Raman *Indian J Met Hydrl Geophys* **29** (1978) 425
5 V Gowarikar, V Thapliyal, S M Kulshrestha, G S Mandal, N Sen Roy and D R Sikka *Mausam* **42** (1991) 25
6 V Gowarikar, V Thapliyal, R P Sarkar, G S Mandal and D R Sikka *Mausam* **40** (1989) 115
7 H N Srivastava and S S Singh *Mausam* **44** (1993) 29
8 H S Bedi and M M S Bindra *Tellus* **32** (1980) 296
9 R N Iyengar *Proc Indian Acad Sci (Earth Planet Sci)* **100** (1991) 105
10 E N Lorenz *Empirical Orthogonal Functions and Statistical Weather Prediction.* Sc. Report I. Statistical Forecasting Project, M.I.T., Deptt. of Meteorology, U.S.A. (1956) 49
11 R W Preisendorfer *Principal Component Analysis in Meteorology and Oceanography* Elsevier Publns, New York (1988)

VARIATIONAL DATA ASSIMILATION WITH AN ADIABATIC VERSION OF THE LMD GCM USING ADJOINT METHOD

H C UPADHYAYA, SANGEETA AGARWAL AND
O TALAGRAND*

*Centre for Atmospheric Science, Indian Institute of Technology
Hauz Khas, New Delhi-*110 016, *India*

(*Received* 14 *June* 1993; *Accepted* 5 *July* 1993)

Variational four-dimensional data assimilation is performed using an adiabatic version of the Laboratoire de Meterologie Dynamique (LMD) multi-level grid point global model. The low resolution model used here has 8 parallels of 16 points each with 5 sigma levels. The cost function or distance dunction consists of a square sum of the difference between the model forecast and obervations over a time interval. The derivation of the adjoint of the LMD GCM (General Circulation Model) is presented. This adjoint model is then used to compute the gradient of the distance function with respect to the initial conditions. This is performed by a single backward time integration of the derived adjoint equations. The optimal solution is obtained by using the conjugate gradient method. The test experiment results show that the variational assimilation technique performed well. The main objective of this study, that is, to develop the adjoint code for the adiabatic version of the LMD GCM is achieved and tested successfully.

Key Words: Adjoint Technique; Conjugate Gradient Method; Cost Function; Distance Function; Descent Methods; Variational Assimilation

Introduction

Numercial modelling has now become a major tool for the study of atmospheric and oceanic dynamics such as weather forecasting and its theoretical and fundamental aspects. Numerical models are computer programs which compute the temporal evolution of the atmosphere and oceanic flow from specified initial and boundary conditions. These conditions control the evolution of the solution in space and time.

The advent of powerful computers has allowed the development of very high resolution models, such as those which are used for operational medium-range weather forecasting. The heterogeneous distribution of data in space and time are available from satellite, and other remote sensing devices. There is a need for an improved assimilation system because of these advances in numerical modelling and the observation network.

There are many techinques used in data assimilation. But considerable attention has been focussed on four-dimensional variational data assimilation[1-]. This powerful technique is also applicable to sensitivity and analysis[8-10] and to parameter estimation. Another application of adjoint equations is the determination of the perturbations which, when superimposed on given meteorological fields, will imply most rapidly.

---

*Laboratoire de Meteorologie Dynamique, Ecole Normale Superieur, 24, rue Lhomond, 75231, cedex 05, Paris (France)

The objective of variational data assimilation is to define,.as accurately as possible, the optimal solution which requires at each iteration the gradients of the cost-function with respect to all the adjusting parameters (e.g., the initial conditions) of a numerical model by fitting the model dynamics to data over a given period of time, where the optimality is measured by a cost-function that expresses the degree of discrepancy or distance function between the model and the data.

Hoffmann's[3] way of obtaining these gradients was found to be impossible to implement on present-day computers. A feasible, yet still expensive, way of evaluating these gradients is through the adjoint technique.

The variational approach to assimilation, with explicit use of the model adjust equations, seems to have been first suggested by Penenko and Obraztsov[13]. They studied a simple linear, small-dimensional example, which nevertheless contained all the basic ingradients of the method. The equation was studied again by Lewis and Derber[2] and La Dimet and Talagrand[1]. Since then, a fairly large number of works have been performed[6,7,14,15]. The gradient of the distance function is computed by first integrating the forecast model forward and then by integrating the adjoint model equations backward in time. The forcing terms are added to the solutions of the adjoint equations at times when observational data are available. Once the values of the distance function and its gradients are available, different minimization methods can be employed to find the optimal solution. In this paper, results are presented for a variational assimilation system being developed for the low-resolution version of the LMD GCM. In the next section, after a brief description of the LMD model and its discretization in space and time, we define the distance function and derive the adjoint equations for the model. Here we also deal with the verification of the correctness of the adjoint code and the gradient. Results of this study are presented in the third section. A summary and general conclusions are discussed in the last section.

Variational Formalism

Brief Description of the Model

The model equations are formualted using cylindrical; coordinates x (longitude) and y ($\sin\phi$, where ϕ is the latitude). The mapping is area-preserving. Horizontal wind velocity is equivalently described by its natural components in the zonal and meridional directions.

The horizontal grid is chosen regular in the (x, y) plane. The density of grid points on the sphere is therefore uniform; however, anisotrophy becomes quite large in the vicinity of both poles. The actual model has 48 parallels of 64 points each. The vertical coordinate is chosen as the normalized pressure $\sigma = p/p_s$. The layers are evenly spaced in an auxiliary coordinate z, the mapping $z \rightarrow \sigma$ being an analytic mapping $(0,1)$ onto itself. The number of layers is 11. In σ coordinate, the layers are unevenly spaced. They are relatively thinner in the vicinity of the ground and near top. The horizontal grid is Arakawa-C. The basic thermodynamic variable in the model is potential enthalpy rather than temperature. In order to define temperature without ambiguity, we introduce an alternative independent vertical variable

$$s = \sigma^k = \pi / \pi_s$$

defined at the layer cores. Values of σ will be considered at interfaces only; the correspondence between discrete values of s and discrete values of σ is given by

$$\overline{\sigma \delta_z s^z} = k s \delta_z \sigma.$$

For more details on the LMD GCM, *see* Sadourny and Laval.

The Adjoint Method

For any continuous linear operator **L** defined on E, it is possible to define its adjoint **L***, which is also a linear operator on E, characterized by the property that, for any tow vectors **X** and **Z** in E the following equality holds between inner products

$$\langle \mathbf{X}, \mathbf{L}\mathbf{Z} \rangle = \langle \mathbf{L}^*, \mathbf{X}, \mathbf{Z} \rangle.$$

In discretized problems, **X** and **Z** are vectors and **L** is a matrix, with **L***, being its transpose. The objective of variational assimilation is to find optimal solution to a given model that will best fit observed fields distributed over some space and time interval. One possible measure of the fit, the distance or cost function, J, consists of weighted squared norm of the difference between the observations at time t and the corresponding components of a field variable.

Let us consider a system whose state at any time t is defined by a state vector $Y(t)$ belonging to some space ε and whose time evolution is described by the equation

$$\mathrm{d}Y/\mathrm{d}t = H(Y), \qquad \ldots (1)$$

where H is regular function of ε into itself. Assuming that observations $Y^0(t_1)$, $Y^0(t_2)$, ..., $Y_0(t_n)$ of the state of the system are available at times $t_1 < t_2 < \ldots < t_n$, we want to find a solution $Y(t)$ of (1) minimizing the functional

$$J[Y(t)] = \sum \langle Y(t_i) - Y^0(t_i), Y(t_i) - Y^0(t_i) \rangle, \qquad \ldots (2)$$

where $\langle , \rangle$ denotes an inner product defined over ε. The observations here are not exactly compatible with a solution of (1), therefore, J is not equal to zero.

Any solution $Y)t)$ of (1) is uniquely defined by the specification of the initial condition $Y(t_1)$. The problem is now to find the initial condition $Y(t_1)$ such that the corresponding solution of (1) minimzes the distance function (2).

Given a solution $Y(t)$ of (1), the first order variation δJ resulting from a perturbation $\delta Y(t_1)$ of the initial condition is equal to

$$\delta J = 2 \sum \langle Y(t_i) - Y^0(t_i), \delta Y(t_i)\rangle , \qquad \dots (3)$$

where the first-order variation $\delta Y(t_i)$ are themselves governed by the linear tangent equation

$$d\delta Y/dt = H'(t)\,\delta Y \qquad \dots (4)$$

where H' is the derivative of H with respect to Y, evaluated at Y. System (4) is obtained by linearising the basic evolution equation (1) about the solution $Y(t)$. For any solution of (4), the value $\delta Y(t)$ at a given time t depends linearly on the initial condition $\delta Y(t_1)$. For simple forward time integration scheme as an example, one can have

$$\delta Y_i = (I + \Delta t H'_{i-1})\delta Y_{i-1} = (I + \Delta t H'_{i-1})(I + \Delta t H'_{i-2})\dots(I + \Delta t H'_1)\,\delta Y_1 \qquad \dots (5)$$

Thus,

$$\delta J = 2 \langle \sum (I + \Delta t H_1'^{*})\dots(I + \Delta t H_{i-1}'^{*})(Y(t_i) - Y^0(t_i)), \delta Y_1\rangle .$$

The summation in this expression is the gradient $\nabla_{Y_1} J$. The ith in this series is obtained by a backward finite-difference integration of the adjoint equation

$$-d\delta^* Y/dt = H'^*(t)\,\delta^* Y$$

from time step i to the initial step starting from

$$\delta^* Y = 2(Y_i - Y_i^0).$$

In summary, the gradient ∇J corresponding to some initial condition $Y(t_1)$ can be explicitly obtained by performing the following operations:

(1) Starting from $Y(t_1)$, integrate the basic equation (1) from t_1 to t_n. Store the solution at each time step Y_i ($i = 1, \dots, n$), which will make up coefficients of the operator $H'^*(t)$.

(2) Starting from $\delta^* Y(t_n) = 2\,Y(t_n) - Y_0(t_n))$, integrate the adjoint equations backward in time from t_n to t_1, the forcing term $2\,Y(t_i) - Y^0(t_i))$ being added to the currently computed solution $\delta^*(Y(t_i)$ at each observation time t_i. The final result at time t_1 is the gradient ∇J.

Examples of deriving H'^* have been given by Talagrand and Courtier[4]. The adjoint code can be constructed directly from the model code in a simple way, which is also effective to easy error detection in adjoint code. Thus, the derivation of the adjoint model in both forms, the differential or the discrete form, is not required.

The Adjoint Model

The GCMs are rather long and complicated codes. The length and complexity of the corresponding adjoint codes is comparable to those of the direct models. One should write adjoints with a few basic principles in mind. The first principle is the principle of locality, i.e., whenever a local modification is made in the direct code, the corresponding modification must also be local in the adjoint code. Locality in turn requires readability, meaning that when a local modification is made in the direct code, it must be easy to locate the place where the corresponding modification is to be made in the adjoint code. This requires transparent notations, between the direct and adjoint codes, as regards subroutine and variable names, instruction labelling, etc.

These requirements for locality and readability clearly imply that the adjoint code must be developed directly from the basic direct code. They also imply that the components of the adjoint code must be in one-to-one correspondence with the components of the direct code: to each subroutine of the direct code, performing a well defined task, will correspond a subroutine of the adjoint code, performing the adjoint task. For example, a typical Fortran statement in the direct code is

$$X = AYZ + BZ^2,$$

where A and B are constants and X, Y, and Z are variables. The corresponding Fortran statement in the tangent linear code is

$$X = A((aZ)Y + (aY)Z) + 2B(aZ)Z,$$

where aY and aZ are the values of Y and Z stored during the forward integration of the model, and Y, and Z now represent the perturbation variables. The corresponding adjoint code will be

$$Z' = Z' + [2B(aZ) + A(aY)]X'$$

$$Y' = Y' + A(aZ)X$$

and $$X' = 0,$$

where X', Y' and Z' represent the gradient of J with respect to the direct code variables X, Y, and Z, respectively (*see* Appendix also). The above form of statements clearly shows that all adjoint variables must be set to zero in the adjoint code before they are used for the first time. The general rules which have just been demonstrated can be very automatically implemented on a Fortran code. It is seen that the tangent linear code is a necessary intermediary for developing the adjoint code. The linear tangent code is also extremely useful for checking the adjoint code. The only basic choice to make, when developing the adjoint code or the linear tangent code, is whether the quantities appearing in nonlinear terms in the direct code will have to be stored at the time of forward time integration, or will have to recalculated when needed. For the present study only the prognostic variables are stored during the forward run.

Verification of the Adjoint Code

The safest way to check the correctness of the adjoint code is to check the adjointness equality $\langle v, Lu\rangle = \langle L^* v, u\rangle$ between scalar products. The best way to check the tangent linear code is to verify that it leads to an exact first-order approximation for the output perturbation, i.e., to verify that the difference $\Delta v - \delta v$, where Δv is the exact, finite difference, output perturbation computed by the tangent linear code, is of order $0(\delta u)$. The same method can be used for checking the exactness of a gradient produced by an adjoint code, i.e.

$$J(u+\alpha\delta u) - J(u) - \langle \alpha\delta u, \nabla J(u)\rangle = 0(\alpha^2), \qquad \ldots (6)$$

For values of α that are small but not too close to the machine zero, one should expect to obtain the required order (Table I) for the above equation (6). All routines of the adjoint code of the LMD GCM were tested both ways.

Table I

Verification of relation (6)

Values of α	Order (α^2)
10^{-1}	0.206908445641E+05
10^{-2}	0.206908835595E+03
10^{-3}	0.206908874577E+01
10^{-4}	0.206908878452E−01
10^{-5}	0.206908878762E−03
10^{-6}	0.206908877412E−05
10^{-7}	0.206908863578E−07
10^{-8}	0.206908975217E−09
10^{-9}	0.206907448229E−11
10^{-10}	0.206908363666E−13
10^{-11}	0.207009968835E−15
10^{-12}	0.207777719943E−17

Distance Function

For the test experiment described in this paper the distance function is defined as

$$J = \sum_{t=1}\sum_{ijl}\{W_q(q_{ijlt} - q^0_{ijlt})^2 + \sum_{t=1}\sum_{ij} W_{ps}(p_{sijl} - p^0_{sijt})^2,$$

where $\sum_{ijl}$ and $\sum_{ij}$ are the summations over all observation points, $\sum_t$ is the summation over time, and W's are weights, and superscript 0 denotes observation, and $q=(u, v, h)$, here h denotes potential enthalpy; p_s is the area-weighted surface pressure, and u and v are covariant vector components of the horizontal wind. The variables, i and j are the horizontal indices, and l is the vertical index

Table II

Convergence of the distance function and the gradient with number of iterations

No. of iterations	normalized distance function (J/J_0)	normalized gradient
0	1	1
1	$1.43*10^{-2}$	$1.2*10^{-1}$
2	$1.15*10^{-5}$	$3.3*10^{-3}$
3	$3.55*10^{-7}$	$5.9*10^{-4}$
4	$5.41*10^{-9}$	$7.1*10^{-5}$
5	$1.86*10^{-10}$	$1.2*10^{-5}$
6	$1.73*10^{-11}$	$3.6*10^{-6}$
7	$9.23*10^{-12}$	$3.0*10^{-6}$
8	$6.12*10^{-13}$	$7.0*10^{-7}$
9	$5.41*10^{-14}$	$2.1*10^{-7}$
10	$3.67*10^{-15}$	$5.4*10^{-8}$
11	$4.28*10^{-16}$	$2.2*10^{-8}$
12	$2.58*10^{-17}$	$4.7*10^{-9}$

The observation is available at all grid points and at each time step of the model. The distance function J is non-dimensional.

Descent Algorithms

Once a numerical algorithm is available for computing ∇J, a descent algorithms can be used in order to determine the value $Y_{\min}$ which minimizes J. Successive estimates Y_n of $Y_{\min}$ are obtained through descent steps of the form

$$Y_{n+1} = Y_n - \rho_n D_n,$$

where, for each n, D_n is a descent direction determined from the successive gradient ∇J_n, ∇J_{n-1}, ..., and ρ_n is an appropriate scalar. If D_n and ρ_n are properly chosen, the sequence Y_n will tend to $Y_{\min}$. Three classical descent algorithm are the steepest descent algorithms, the conjugate gradient algorithm and the quasi-Newton, or variable metric algorithm. These algorithms are described, and their properties compared, in Gill *et al.*

Test Experiment and Result

As mentioned earlier that all routines of the adjoint code have been tested separately using linear tangent code and by computing the gradients. Besides this the whole system (the direct code, the adjoint code and the minimization procedure) was tested. The forecast model was run for only 5 steps from a specified initial condition and at each time step the field variables (only prognostic ones) were stored as observations. Starting from zero initial conditions for the adjoint model while the ∇J were inserted whenever an analysis time was reached and this recovered the original initial field through the iterative procedure: that is, the distance function as well as the gradient should reduce to zero.

The test run was conducted with FGGE data with low resolution. The run generates the observation data at each time step. The adjoint integration was started from final step to initial time. The conjugate gradient method was employed to achieve optimal solution. Table II shows the convergence of the distance function and the gradient with number of iterations. The distance function and the gradient have reduced to several orders of magnitude in only 12 iterations.

Summary

We have completed writing the adjoint of the dynamics part of the LMD GCM. The correctness of the adjoint model has been tested by the methods discussed in this paper. The same weight is used for each variable at each level. Model generated data is used as observation for the present study. Conjugate gradient method is used to attain the optimal solution of the test experiment. The test run with realistic initial conditions seems to be satisfactory and encouraging. A reasonable reduction in the distance function was achieved, and the reproduction of initial field is quite satisfactory (the root mean square of the field variables is the order of 10^{-6} units). Future work involves the addition of 'physics' in the adjoint code and perform some sensitivity experiments using a moderate resolution GCM.

Acknowledgements

The first author is grateful to Professor R Sadourny, Director LMD for providing an opportunity to work at his Laboratory, and to the Department of Science and Technology, Government of India. This work is undertaken under the Indo-French Centre for Promotion of Advanced Research/Centre Franco-Indian Pour la Promotion de Recherche Avancée (Project No. 711-1).

References

1 F X Le Dimet, O Talagrand *Tellus* **38A** (1986) 97-110
2 J M Lewis and J C Derber *Tellus* **39A** (1985) 309-322
3 R N Hoffmann *Mon Wea Rev* **114** (1980) 388-397
4 O Talagrand and P Courtier *Q J R met Soc* **113** (1987) 1311-1328
5 P Courtier and O Talagrand *Q J R met Soc* **113** (1987) 1329-1347
6 Winston C Chao and Chan Lang-Ping *Mon Wea Rev* **120** (1992) 1661-1673
7 I M Navon, X Zou, J Derber and J Sela *Mon Wea Rev* **120** (1992) 1433-1446
8 D G Cacuci *J math Phys* **35A** (1981) 2794-2802
9 M C G Hall and D G Cacuci *J atmos Sci* **40** (1983) 2537-2546
10 M C G Hall, D G Cacuci and M E Schlesinger *J atmos Sci* **30** (1982) 2038-2050
11 B Urban *Maximum Error Growth in Simple Meteorological Models Tech* Rep Ecole Normale de la Meteorologie, Toulouse (1985) (French)
12 J F Lacarra and O Talagrand *Tellus* **40A** (1988) 81-95
13 V V Penenko and N N Obraztsov *Sov met Hydrol* **11** (1976) (*Trans. Russian*)
14 A C Lorenc *Q J R met Soc* **114** (1988) 205-240
15 P Courtier and O Talagrand *Tellus* **42A** (1990) 531-549
16 R Sadourny and K Laval *New Perspectives in Climate Modelling* (*Eds.* N Berger and C Nicolis) No 16 In *Elsevier Sci* (1984) 173-197
17 P E Gill, W Murray and M H Wright *Practical Optimizatioin* Academic Press (1981) 401 pp

Appendix

We present here one of the several routines used in this study. In the DIRECT CODE, pbarx, for example, is a variable. The same is perturbation in the LINEAR TANGENT CODE, and in the ADJOINT CODE it is a gradient. The variable apbarx is the value of pbarx stored during the forward integration of the model.

```
*DIRECT CODE or FORWARD CODE***
```

```
SUBROUTINE flumass (pbarx, pbary, vcont, ucont, pbaru, pbarv)
*CALL paramet
DIMENSION pbarx (iplijmpl), pbary (ipljm), vcont (ipljm, max),
*ucont (ipljmpl, max), pbaru (ipljmpl, max), pbarv (ipljm, max)
*CALL const
COMMON/scratch/apbarun (iipl), apbarus (iipl), aire4n (iipl),
*aire4s (iipl)
EXTERNAL SSUM
DO 5l = l, max
DO 1 jij = iip2, ipljm
pbaru (ij, l) = pbarx (ij)*ucont(ij, l)
1 CONTINUE
DO 3 ij = l, ipljm
pbarv (ij, l) = pbary (ij)*vcont (ij, l)
3 CONTINUE
5 CONTINUE
sairen = SSUM (iim, aire (l), l)
saireun = SSUM (iim, aireu (l), l)
saires = SSUM (iim, aire (ipljm + 1), l)
saireus = SSUM (iim, aireu (ipljm + 1), l)
DO 20 l = 1, max
ctn = SSUM (iim, pbarv (1,l), 1)/sairen
cts = SSUM (iim, pbarv (ipljmil + 1, l), 1)/saires
pbaru (1, l) = pbarv (1, l)-ctn*aire(1)
pbaru (ipljm + 1, l) = – pbarv (ipljmil + 1, l) + cts*aire (ipljm + 1)
DO 11 i = 2, iim
pbaru (i, l) = pbaru (i – 1, l).+
pbarv (i, l) – ctn*aire(i)
pbaru (i + ipljm, l) = pbaru (i + ipljm – 1, l)
pbarv (i + ipljmil, 1) + cts*aire (i + ipljm)
11 CONTINUE
DO 12i = 1, iim
apbarun (i) = aireu (i)*pbaru (i, l)
apbarus (i) = aireu (i + ipljm)*pbaru (i + ipljm, l)
12 CONTINUE
ctn0 = – SSUM (iim, apbarun, 1)/saireum
cts0 = -SSUM (iim.apbarus, 1)/saireus
DO 14i = 1, iim
pbaru (i, 1) = 2, *(pbaru(i, l) + ctn0)
pbaru (i + ipljm, l) = 2, *(pbaru (i + ipljm, l) + cts0)
```

```
14 CONTINUE
pbaru (iipl, l) = pbaru (1, l)
pbaru (ipljmpl, 1) = pbaru (ipljm + 1, l)
20 CONTINUE
RETURN
END
```

```
***LINEAR TANGENT CODE*

SUBROUTINE tflumass (pbarx, pbary, vcont, pbaru, pbarv,
*apbarx, apbary, avcont, aucont, apbaru, apbarv)
*CALL paramet
DIMENSION pbarx (ipljmpl), pbary (ipljm), vcont (ipljm, max),
*ucont (ipljmpl, max), pbaru (ipljmpl, max), pbarv (ipljm, max)
DIMENSION apbarx (ipljmpl), apbary (ipljm), avcont (ipljm, max),
*aucont (ipljmpl, max), apbaru (ipljmpl, max) apbarv (ipljm, max)
*CALL const
COMMON/safluma/apbarun (iipl), apbarus (iipl), aire4n (iipn),
*aire4s (iipl)
EXTERNAL SSUM
DO 5l = 1, max
DO 1 ij = iip2, ipljm
pbaru (ij, l) = pbarx (ij) *aucont(ij, l) + apbarx (ij) *ucont (ij, l)
1 CONTINUE
DO 3 ij = 1, ipljm
pbarv (ij, l) = pbary (ij) *avcont (ij, l) + apbary (ij) *vcont (ij, l)
3 CONTINUE
5 CONTINUE
sairen = SSUM (iim, aire (1), l)
saireun = SSUM (iim, aireu (1), l)
saires = SSUM (iim, aire (ipljm + 1), l)
saireus = SSUM (iim, aireu (ipljm + 1), l)
DO 20l = 1, max
ctn = SSUM (iim, pbarv (1, l), l)/ sairen
cts = SSUM (iim, pbarv (ipljmil − 1, l), l)/saires
pbaru (1, l) = pbarv (1, l) − ctn*aire (1)
pbaru (ipljm + 1, l) = − pbarv (ipljmil + 1, l) + cts *aire (ipljm + 1)
DO 11i =2, iim
pbaru (i, l) = pbaru (i − 1, l) +
pbarv (i, l) − ctn *aire (i)
pbaru (i + ipljm, l) = pbaru (i + ipljm-1, l),
pbarv (i + ipljmil, l) + cts *aire (i + ipljm)
11 CONTINUE
DO 12i = 1, iim
abparun (i) = aireu (I) *pbaru (i, l)
apbarus (i) = aireu (i + ipljm) *pbaru (i + ipljm, l)
12 CONTINUE
ctn0 = − SSUM (iim, apbarun, l)/saireun
cts0 = − SSUM (iim, apbarus, l)/saireus
```

```
DO 14i = 1, iim
pbaru (i, l) = 2, *(pbaru (i, l) + ctn0)
pbaru (i + ipljm, l) = 2, *(pbaru (i + ipljm, l) + cts0)
14 CONTINUE
pbaru (iipl, l) = pbaru (1, l)
pbaru (ipljmpl, l) = pbaru (ipljm + 1, l)
20 CONTINUE
RETURN
END
```

ADJOINT CODE

```
SUBROUTINE aflumass (pbary, vcont, ucont, pbaru, pbarv,
*apbarx, apbary, avcont, aucont, apbaru, apbarv)
*CALL paramet
DIMENSION pbarx (ipljmpl), pbary (ipljm), vcont (ipljm, max),
*ucont (ipljmpl, max). pbaru (ipljmpl, max), pbarv (ipljm. max)
DIMENSION apbarx (ipljmpl), apbary (ipljm), avcont (ipljm, max),
*aucont (ipljmpl  max), apbaru (ipljmpl, max), apbarv (ipljm, max)
*CALL const
COMMON/safluma/apbarun (iipl), apbarus (iipl), aire4n (iipl),
*airc4s (iipl)
EXTERNAL SSUM
DO 45i = 1, iipl
apbarun (i) = 0.
45 apbarus (i) = 0.
sairen = SSUM (iim, aire (1), l)
saireun = SSUM (iim, aireu (1), l)
saires = SSUM (iim, aire (ipljm + 1), l)
saireus = SSUM (iim, aireu (ipljm + 1), l)
cts0 = 0.
ctn0 = 0.
DO 20l = max, 1, - 1
pbaru (1, l) = pbaru (iipl, l)
pbaru (ipljm + 1, l) = pbaru (ipljmpl, l)
DO 14i = iim, 1, - 1
cts0 = cts0 + 2, *pbaru (i + ipljm, l)
ctn0 = ctn0 + 2, *pbaru (i, l)
pbaru (i + ipljm, l) = 2, *pbaru (i + ipljm, l)
pbaru (i, l) = 2, *pbaru (i, l)
14 CONTINUE
DO 46i = 1, iim
apbarus (i) = apbarus (i)-cts0/saireus
apbarun (i) = apbarun (i)-.ctn0/sairen
46 CONTINUE
cts0 = 0.
ctn0 = 0.
DO 12i = iim, 1, -1
pbaru (i + ipljm, l) = pbaru (i + ipljm, l) + aireu (i + ipljm) *apbarus (i)
pbaru (i, l) = pbaru (i, l) + aireu(i) *apbarun (i)
apbarun (i) = 0.
apbarus (i) = 0.
12 CONTINUE
cts = 0.
cts = 0.
DO 11i = iim, 2, - 1
cts = cts + aire (i + ipljm) *pbaru (i + ipljm, l)
pbarv (i + ipljmil, l) = pbarv (i + ipljmil, l) - pbaru (i + ipljm, l)
pbaru (i + ipljm-1, l) = pbaru (i + ipljm-1, l) + pbaru (i + ipljm, l)
pbaru (i + ipljm, l) = 0.
ctn = ctn-aire (i)*pbaru (i, l)
pbarv (i, l) = pbarv (i, l) + pbaru (i, l)
pbaru (i-1, l) = pbaru (i-1, l)+ pbaru (i, l)
pbaru (i, l) = 0.
11 CONTINUE
cts = cts + aire (ipljm + 1) *pbaru(ipljm + 1, l)
pbarv (ipljml + 1, l) = pbarv(ipljmil + 1, l)-pbaru (ipljm + 1, l)
pbaru (ipljm + 1, l) = 0
ctn = ctn-aire (1) *pbaru (1, l)
pbarv (1, l) = pbarv (1, l) + pbaru (1, l)
pbaru (1, l) = 0.
DO 21i = iim, 1, - 1
pbarv (ipljmil + i, l) = pbarv (ipljmil + i, l) + cts/saires
pbarv (i, l) = pbarv (i, l) + ctn/sairen
21 CONTINUE
cts = 0.
ctn = 0.
20 CONTINUE
DO51 l=max, 1, - 1
DO 3 ij = 1, ipljm
vcont (ij, l) = vcont (ij, l) + apbary (ij)*pbarv (ij, l)
pbary (ij) = pbary (ij) + avcont (ij, l) *pbarv (ij, l)
pbarv (ij, l) = 0.
3 CONTINUE
DO 1 ij = iip 2, ipljm
ucont (ij, l) = ucont (ij, l) + apbarx (ij) *pbaru (ij, l)
pbarx (ij) = pbarx (ij) + aucont (ij, l) *pbaru (ij, l)
pbaru (ij, l) = 0.
1 CONTINUE
5 CONTINUE
RETURN
END
```

DYNAMICAL MODELS FOR QUANTITATIVE PRECIPITATION FORECASTS IN INDIA

P N SEN
*India Meteorological Department, Shivajinagar, Pune-*411005, *India*

(*Received* 16 *August* 1993; *Accepted* 19 *August* 1993)

For efficient flood forecasting, Quantitative Precipitation Forecast (QPF) is the most important input. There exist several models for QPF. Recently developed dynamical models for QPF have been highlighted. In India, there are essentially two dynamical models for QPF. A detailed discussion on these models is presented and the results obtained are critically analysed.

Key Words: Flood Forecasting; Dynamical Model; Precipitation Efficiency; Precipitation Rate; Quantitative Precipitation Forecast; Vertical Velocity

Introduction

Floods, especially flash floods rank high in the list of natural disasters for most countries of the world. Accurate prediction of precipitation or precipitation rate is crucial to the prediction of floods in real time. Attempts have been made in several countries to forecast the amount of precipitation in different catchment areas with different lead time (the difference between the time of occurrence of the forecasted phenomenon and the time when the forecast is issued) using various techniques but achieved only limited success. A very good account on the status of the Quantitative Precipitation Forecast (QPF) models used in Australia, Canada, France, Japan, Germany, Sweden, Switzerland, USSR, and USA for operational purposes has been given by Bellocq[1]. Although there exist several models and procedures for the real time Quantitative Precipitation Prediction, only recently dynamic precipitation models suitable for use with hydrologic models in the real time forecasting of floods and flash floods have been developed[2,3]. The recent advances in QPF research and possible future direction towards achieving improved use of QPF information in hydrological forecasting have been discussed at length by Georgakakos and Hudlow[4].

The main rain producing system in the Indian subcontinent is the 'Monsoon'. India is blessed with two Monsoons: Southwest Monsoon (June to September) and Northeast Monsoon (October to December). The Southwest Monsoon accounts for about 80% of the Annual precipitation of the most parts of the country whereas the Northeast Monsoon season is the major rain giving season for the southern peninsula particularly Tamil Nadu. Some parts or other in India come under the grip of flood due to heavy rainfall in the catchment areas every year during the monsoon season causing loss of lives, livestock and damage to standing crops and property. In order to save lives and property timely forecast of the amount of rainfall is essential. India Meteorological Department has been issuing QPF in real time for various river catchments in India from the Flood Meteorological offices located at various locations using different techniques viz., Synoptic, Statistical, Synoptic-Statistical for the past two decades. These techniques are semi-quantitative rather than quantitative and are subjective in nature. The experience of the forecasters with weather patterns and conditions plays a very important role. Of late, some dynamical models have been developed for QPF in India[5,6].

The present paper highlights some of the recently developed Dynamical Models for QPF and the details of the models developed for QPF in India.

Recent Development of Dynamical Models for QPF

The general category of methods dealing with the use of numerical meteorological models which are physically and dynamically based are known as Dynamical Models. These models range in scale from the very large ones to the very small ones consisting of one dimensional microphysical cloud-physics models.

In the United States of America, the Limited Area Fine Mesh (LFM) model and Moveable Fine Mesh (MFM) models[7,8] are used for QPF. These models generally account for a large scale precipitation process as well as a subgrid convective precipitation process. European countries also use NWP models for operational QPF. Some of the countries use Statistical-Dynamical (or Stochastic-Dynamical) or Synoptic-Dynamical models for operational QPF. Even in USA QPF's are not all derived strictly from Computer models and methods. Meteorologists use their hydrometeorological observations, climatology, knowledge of local influences such as orography and current and past weather patterns as well as outputs from the models combined with the meteorologists' own interpretive skills. This type of Man-Machine approach is in vogue in many countries in Weather Forecasting, particularly in QPF. The basic practice used in all the NWP models for QPF is moisture balance coupled with the dynamical equations and inputs are generally temperature, dewpoint, surface pressure, geopotential field, winds etc.

The Stochastic-Dynamic models developed by Georgakakos *et al.*[2,3,9] during the last decade have been successfully used with hydrological models to form integrated hydrometeorological models for real time site-specific flood and flash flood prediction. Initially Georgakakos and Bras[2,3] formulated a one-dimensional precipitation model. Based on the surface pressure, temperature, dew point temperature, their model gives as an output the precipitation rate. Their model state is the mass of the condensed liquid water equivalent in the area characterized by the input temperature and pressure indicies. The model formulation is based on pseudoadibatic ascent of the air masses and on simplified cloud microphysics with exponential distribution of particle size and linear dependence of the particle terminal fall-velocity on the particle diameter. Evaporation of the falling particles, for unsaturated sub-cloud layer is explicitly taken into account by the model.

This formulation has been extended by Georgakakos and Sen[10] so that the model can accept both surface and upper air data as input. In this model, the cloud depth is discretized into several layers where boundary levels are the levels of the upper air data. The condensation inflow rate to the cloud is computed from the layer-average updraft velocity and condensed equivalent mass in each of these layers. The cloud top is determined as the level where pseudo-adiabatic cloud temperature becomes equal to the environmental one available from the upper air data. This formulation effectively eliminates the problem of overprediction in case of temperature inversions present when only surface data are used. It also provides a more sound basis for the definition of the updraft velocity throughout the cloud column and of the cloud top pressure level.

Lee and Georgakakos[11] suggested a two-dimensional model for QPF at spatial and temporal scales relevant to hydrological forecasting. The model uses as input operation forecasts of surface pressure, air temperature, and dewpoint temperature and wind speed and direction at mid tropospheric levels. The two dimensional precipitation model produces as an output mean areal precipitation at scales down to $100 \mathrm{km}^2$ in space and 1hr in time. Thermodynamcis and microphysics are used to determine the source and sink terms in the conservation equation for the mass of condensed liquid water equivalent utilized by the model.

The results obtained by using these Stochastic-Dynamic models are quite encouraging.

Dynamical Models developed for QPF in India

In India, a dynamical model for QPF was first suggested by Upadhyay *et al.*[5]. They took into consideration two parameters for the QPF model (*i*) the distribution of moisture in the air column; and (*ii*) the distribution of vertical velocity at different levels. The main consideration is that a saturated air column of thickness $\mathrm{d}p$ standing over unit area is ascending with a vertical velocity $\omega = \mathrm{d}p/\mathrm{d}t$. The rate of precipitation (R) released by this column varies as $\omega \mathrm{d}p$ or

$$R = C\omega \mathrm{d}p.$$

The Coefficient C is a function of temperature, pressure and moisture content of the layer as it represents rate of precipitation released by the layer of unit thickness (1hPa) ascending with unit velocity ($-$ 1hPa s^{-1}).

The first law of thermodynamics gives

$$\mathrm{d}Q = c_p \mathrm{d}T + g \mathrm{d}z.$$

If $\mathrm{d}q_s$ is the change in saturated specific humidity in the time interval $\mathrm{d}t$, the latent heat of condensation released will be $-L\,\mathrm{d}q_s$

$$-L\,\mathrm{d}q_s = c_p \mathrm{d}T + g \mathrm{d}z.$$

The rate of precipitation, $R = -\dfrac{1}{g}\displaystyle\int \frac{\mathrm{d}q_s}{\mathrm{d}t}\mathrm{d}p.$

Using the above thermodynamic relation the rate of precipitation can be found as

$$R=\frac{c_p}{g}\int\frac{\gamma_d-\gamma_s}{L}\,w\,\mathrm{d}p.$$

Taking $\omega=1$ hPas^{-1}, where $\omega=-g\rho w$, C is found out as

$$C=\frac{c_p}{g}\int\frac{\gamma_d-\gamma_s}{L}\frac{1}{-g\rho}\,\mathrm{d}p,$$

where $\gamma_s=\frac{g}{c_p}(1+Lq_s/R_d T)/(1+\varepsilon L^2 q_s/\mathrm{c_p}\mathrm{R_v}T^2)$.

For the computation of vertical velocity at the top of the frictional layer the formulation suggested by Das[12] was used

$$\omega_0=-\frac{q}{f^2}\rho_0\Big/\sqrt{\frac{Kf}{2}}\sin(2\delta)\cdot\nabla^2\phi,$$

where K is the Coefficient of eddy viscosity and δ represents the angle between the actual and geostrophic winds, the other terms have their usual meaning.

The model was used for QPF in the case of a depression over central India. The rainfall forecast issued using the above formulation was in fair agreement with actual rainfall distribution (cf: Figs 2&3 of [5]).

The main drawback of this model is that advection of moisture has not been considered. This was overcome in the model suggested by Sen[6]. He developed a dynamical model for QPF based on physical and dynamical laws.

His dynamical model is based on the hypothesis that the specific humidity in an air column is conserved. In other words, it is based on the principle of conservation of specific humidity. In mathematical notation it can be expressed as:

$$\frac{\mathrm{d}q}{\mathrm{d}t}=\frac{\partial q}{\partial t}+u\frac{\partial q}{\partial x}+v\frac{\partial q}{\partial y}+\omega\frac{\partial q}{\partial p}=0, \qquad \ldots(1)$$

where $u=\frac{\mathrm{d}x}{\mathrm{d}t}$, $v=\frac{\mathrm{d}y}{\mathrm{d}t}$ and $\omega=\frac{\mathrm{d}p}{\mathrm{d}t}$.

The precipitable water vapour in a column of air is the total mass, R_F, of water vapour per unit area in the column. Symbolically[13],

$$R_F=\int_{z_s}^{z_t}\rho_v\,\mathrm{d}z, \qquad \ldots(2)$$

where ρ_v is the density of water vapour and z_s and z_t are the elevations of the bottom and the top of the air column respectively.

The relation (2) can be written in the isobaric coordinate as

$$R_F = -\int_{p_s}^{p_t} \frac{\rho v}{g\rho}\,\mathrm{d}p, \qquad \ldots(3)$$

where ρ is the density of air and p_s and p_t the pressure at the bottom and at the top of the air column respectively and g, the acceleration due to gravity.

Eq. (3) can be written as

$$R_F = -\frac{1}{g}\int_{p_s}^{p_t} q\,\mathrm{d}p, \qquad \ldots(4)$$

where $q = \rho_v/\rho$ = Specific Humidity.

Since q is a nondimensional quantity, the variable R_F has the dimensions of mass per unit area. R_F may be converted to the parameter "Precipitable Water" by dividing it by the density of water.

An expression for the Rate of Precipitation, can be derived from equation (4) in the following way

$$\frac{\partial R}{\partial t} = E\frac{\partial R}{\partial t}F, \qquad \ldots(5)$$

where E $(0 \leqslant E \leqslant 1)$ is a multiplication factor. The multiplication factor has been used because R_F is the amount of precipitable water vapour and not the actual amount of precipitation. $R_F = R$, if and only if all the available moisture condenses and falls as precipitation. But in the actual atmosphere it has been found that only a part of the precipitable water vapour gets converted into precipitation. Thus E is a measure of the proportion of the available moisture which precipitates and may be termed as Precipitation Efficiency. This is a key parameter which takes care of moisture loss due to evaporation and other aspects of cloud microphysics.

Combining equations (4) and (5), the following expression can be obtained

$$g\frac{\partial R}{\partial t} = gE\frac{\partial R_F}{\partial t} = -E\frac{\partial}{\partial t}\Bigg|\int_{n}^{p_t} q\,\mathrm{d}p = -E\Bigg|\int_{n}^{p_t}\frac{\partial q}{\partial t}\,\mathrm{d}p. \qquad \ldots(6)$$

It has been assumed in the above that the variation of p_t and p_s with time counterbalance each other.

$\because$ Eqs. (1) and (6) give

$$g\frac{\partial R}{\partial t} = E\int_{p_s}^{p_t}\left(u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y}\right)\mathrm{d}p + E\int_{p_s}^{p_t}\omega\frac{\partial q}{\partial p}\,\mathrm{d}p$$

or;

$$g\frac{\partial R}{\partial t} = E\sum_{i=1}^{n}(p_i+1-p_i)\left\langle\frac{u\partial q}{\partial x}+\frac{v\partial q}{\partial y}\right\rangle i, i+1^{+}E\sum_{i=1}^{n}(q_i+1-q_i)\langle\omega\rangle_{i,i+1} \qquad \ldots(7)$$

where $i=1$, corresponds to the surface pressure p_s and $i=n$ to the pressure surface p_t and $\langle\ \rangle_{i,i+1}$ represents average value of the layer bounded by the surfaces i and $i+1$.

Eq. (7) can also be written as

$$\frac{\partial R}{\partial t} = E\sum_{i=1}^{n}(\Delta p)_{i,i+1}\langle\mathbf{V}\cdot\nabla q\rangle_{i,i+1} + E\sum_{i=1}^{n}(\Delta q)_{i,i+1}\langle\omega\rangle_{i,i+1}, \quad \ldots (8)$$

where $\quad (\Delta p)_{i,i+1} = p_{i+1} - p_i; (\nabla q)_{i,i+1} = q_{i+1} - q_i$

The first term on the right-hand side of eq. (8) actually represents the horizontal advection of moisture and may be termed as the Advection Term and the second term represents the effect of the vertical velocity on moisture and may be termed as the Vertical Velocity Term. Thus eq. (8) shows that the Velocity field especially the vertical velocity coupled with information on humidity, is related to the rate of precipitation in large area.

Precipitation is the end product of the physical processes taking place in the atmosphere. The occurrence of precipitation is strongly controlled by the motion of the cloud air. In other words, rainfall is always associated with clouds and the moisture advection normally takes place at the boundary layer which almost coincides with the base of the clouds. It would be quite reasonable if 850hPa level is chosen as the top of the boundary layer. In that case the summation in the first term on the right-hand side of eq. (8) need not be performed up to the top of the air column; the summation up to the top of the boundary layer would be sufficient. Thus the eq. (8) may be modified as

$$g\frac{\partial R}{\partial t} \approx E(\Delta p)_{p_s}, 850\,\langle\mathbf{V}\cdot\nabla q\rangle_{p_s}, 850 + E\sum_{n=1}(\Delta q)_{i,i+1}\langle\omega\rangle_{i,i+1} \quad \ldots (9)$$

Since the parameter q, the specific humidity, is not directly measured it is desirable that it be expressed in terms of some parameters, which are either measured at the observational sites or at least reported in the synoptic or upper air observations.

By definition, the specific humidity is given by

$q = \rho_v/\rho = \rho_v/(\rho_d + \rho_v)$, ρ_d = density of the dry air

$$= \frac{0.622\, e_s(T_d)}{p - 0.378\, e_s(T_d)}, \quad \ldots (10)$$

where $e_s(T_d)$ is the saturation vapor pressure over a plane surface of pure water; T_d; the dew point temperature. The saturation vapor pressure is a non-linear convex function of temperature. A convenient formulation for determining the saturation vapor pressure $e_s(T_d)$ in terms of T_d may be used for this purpose. The polynomial relation suggested by Lowe and Ficks[14] has been found to provide an excellent fit with the observed ones in the range −50°C to +50°C (Pruppacher and Klett[15]).

The formulation reads as

$$e_s(T_d)=\sum_{n=0}^{6} a_n T_d^n, \qquad \text{with } T_d \text{ in °C and } \rho_s \text{ in hPa,} \qquad \ldots(11)$$

where $a_0=6.107799961$ $\quad a_1=4.436518521\times10^{-1}$

$a_2=1.428945805\times10^{-2}$ $\quad a_3=2.650648471\times10^{-4}$

$a_4=3.031240396\times10^{-6}$ $\quad a_5=2.034080948\times10^{-8}$

$a_6=6.136820929\times10^{-11}$

Combining eqs. (10) and (11) an expression for q can be obtained as:

$$q=\frac{0.622}{\,}\sum^{6} a_n T_d^n, \qquad \ldots(12)$$

where T_d is expressed in °C and p in hPa and q is a nondimensional quantity.

Using eqs. (9) and (12), the rate of precipitation may be computed provided the value of E is available. Normally the rate of precipitation is expressed in $\text{kgm}^{-2}\ \text{sec}^{-1}$. But since 1 kg m^{-2} of liquid water is equivalent to the depth of 1mm of rainfall $\frac{\partial R}{\partial t}$ can be expressed directly as mm sec^{-1}. The vertical p-velocity ω may be computed using either the equation of continuity in isobaric coordinates or the diagnostic omega equation. The second method is more accurate.

The method suggested by Sulakvelidze[16] and applied by Georgakakos and Bras[2,3] for vertical velocity can also be used. The expression for the vertical velocity w used by them is

$$w=a\sqrt{c_p\Delta T}, \qquad \ldots(13)$$

where a is a constant parameter, $\Delta T=|T_m-T_s|$, T_m is the parcel temperature (°K) at a certain level p (hPa) assuming pseudoadiabatic ascent and T_s is the corresponding ambient air temperature (°K). The square of the quantity 'a' is analogous to the ratio of the Kinetic to the Thermal Energy per unit mass of ascending air, at the level p. Therefore, 'a' is a non-dimensional quantity. When the parcel temperature is more than the environmental temperature there would be upward vertical motion and opposite is the case when the parcel is cooler than the environment. Georgakakos and Bras[2,3] have found out that a value of 0.002 for 'a' gave a good fit. From the eq. (13) the vertical p-velocity can be computed using the following relation (Holton[17]).

$$\omega\simeq-g\rho w. \qquad \ldots(14)$$

The quantity Precipitation Efficiency, E, is a complex, elusive factor to determine quantitatively. It is not theoretically spatially constant either. Rhea[18] used precipitation efficiency (which is slightly different from the precipitation efficiency defined here) in his Orographic Precipitation Model and reported that many computations of precipitation rates for hydrometeorological purposes routinely set $E=1$ (i.e., they equate condensation supply rate to precipitation rate). A large number of test cases were run by Rhea using a variety of

E values and found that on 50% occasions a value of 0.25 for E and at least 70% occasions a value of 0.21 or greater gave a good fit with the observed values. Moreover, it has been estimated that only about 30% of the moisture falls out as precipitation (Haltiner and Martin[13]). Thus it would be reasonable if a value of 0.20 to 0.30 for E is chosen considering the loss due to evaporation of the droplets of precipitation. But it is recommended that the value of E be estimated for each individual station with a long series of data with varieties of rain storms. However, it is to be remembered that extreme parameter sensitivity is not desirable considering the crudeness of the input data as well as the precipitation measurement.

To test the model computation of rainfall rate has been done for New Delhi during the southwest monsoon of 1984 and for Bombay, Ahmedabad and Nagpur for 1989. Upper air data for New Delhi, Bombay, Ahmedabad and Nagpur upto 300hPa level and upper air data along with wind observations for neighbouring stations upto 850hPa levels for 00 UTC have been utilised. The vertical velocity has been calculated using the scheme of Sulakvelidze[16]. The result are presented in the Tables I, II, III and IV.

Table I

Precipitation rate for New Delhi

Date (1984)	Forecast rate of precipitation for the next 24 hours (mm/hr)	Actual rate of precipitation (mm/hr)
1 July	0.94	0.91
2 July	0.62	0.85
3 July	0.25	0.06
4 July	0.00	Trace*
25 Aug	0.40	0.62
26 Aug	2.04	1.71
27 Aug	0.70	1.96
28 Aug	0.08	0.08
29 Aug	0.60	Trace*

*Trace means rainfall of 0.4mm or less in 24 hours.

The model forecasts for rainfal rate are in reasonable agreement with the actual values. In some cases the difference is quite large. However, the correlation between the model forecast and the actual has been found to be 0.60. In the above the vertical p-velocity has been computed using Sulakvalidze scheme. In this scheme the vertical velocity is a function of stability. In other words, if the atmosphere is unstable the vertical velocity will be upward. In case of stratiform cloud which is seen during the southwest monsoon period the vertical velocity computed by the above scheme would be lower than the actual vertical velocity. Thus the model will have the tendency to underestim-

Table II

Precipitation rate for Bombay

Date (1989)	Forecast rate of precipitation for the next 24 hours (mm/hr)	Actual rate of precipitation (mm/hr)	
		Colaba	Santacruz
20 July	0.60	0.06	0.30
21 July	3.02	3.23	3.92
22 July	1.75	2.25	2.91
23 July #	1.03	5.40	4.38
24 July #	1.61	1.01	2.41
25 July	0.42	0.64	0.82
26 July	0.58	0.36	0.53

Based on Radiosonde data of 1200UTC because 0000UTC flight terminated in the lower troposphere and the actual precipitation rate is calculated from the rainfall between 1200UTC and 2400UTC.

Table III

Precipitation rate for Ahmedabad

Date (1989)	Forecast rate of precipitation for the next 24 hours (mm/hr)	Actual rate of precipitation (mm/hr)
23 July	1.10	2.10
24 July	ND	1.41
25 July	0.90	0.34
26 July	1.50**	0.25
27 July	0.00	0.02
28 July	0.80	0.52
29 July	0.02	0.01

ND—No Data;
**Radiosonde data available only upto 700 hPa.

ate the rainfall from stratiform cloud. In the absence of diagnostic Omega values the Sulakvelidze method is the most suitable.

The model developed by Sen[6] has been utilized by Raj *et al.*[19] with some modification for QPF over Thiruvananthapuram and Madras for the Southwest Monsoon and the Northeast Monsoon period of 1992. The Precipitable Water Content (PWC) was computed for every day. The PWC that is likely to be added in the next 24 hours (PWC 24) was also computed. Their computation was only based on the vertical velocity term. The advection term has been neglected, in their computation. The approximate profile of vertical velocity was

Table IV

Precipitation rate for Nagpur

Date (1989)	Forecast rate of precipitation for the next 24 hours (mm/hr)	Actual rate of precipitation (mm/hr)
22 Aug	0.80	0.02
23 Aug	0.30	0.11
24 Aug	0.60	2.19
25 Aug	0.90	0.20
26 Aug	0.30	0.23
27 Aug	0.20	0.24

computed by the slice method based on the radiosonde data. The Saturated PWC (SPWC) and another parameter called percentage of saturation required for condensation (PSRC) were worked out. Based on PWC, PWC 24, SPWC and PSRC a quantity of moisture available for condensation in the next 24 hours (MAC 24) was evaluated. Utilizing the past rainfall data and the MAC 24 of one week, precipitation efficiency was computed for each day and based on this daily rainfall forecast was obtained. The results obtained have been critically analysed. By and large results obtained appeared satisfactory. The results of the forecast have also been evaluated on the basis of a qualitative definition of rainfall amounts by classifying the forecast and realised amounts as dry, very light rain, light rain, moderate rain, heavy rain, very heavy rain etc. A forecast within one class of the realised rain was deemed as correct. According to this methodology nearly 65% and 75% correct forecasts were realised for Thiruvananthapuram and Madras areas respectively.

The model suggested by Sen[6] is very simple and can be used for operational purposes. The model is based on the hypothesis that the specific humidity of the atmospheric column under consideration is conserved. This assumption is valid as long as there is no change in phase which means that once the process of condensation starts the specific humidity may not remain constant. Since in this model, the rate of change of precipitable water vapour in the atmospheric column is calculated first and the rate of precipitation is computed from it through a factor called precipitation efficiency, the above assumption may be considered valid. The same hypothesis has been taken into consideration in the Limited Area Fine Mesh Model (LFM) in the National Weather Service (NWS) of the United States of America[3,8]. Several workers in USSR attempted to develop a method to forecast cloudiness and precipitation on the basis of same hypothesis (for complete list See Matveev[20a]).

Discussion

The model proposed by Sen may be reasonably successful for predicting precipitation associated with large scale (synoptic scale) disturbances and may not be successful to that extent for the prediction of orographic and convective precipitation. The orographic vertical motion is frequently an order of magnitude larger than that associated with large scale vertical velocity, while the vertical motion associated with mesoscale convective system may be still higher. Both orographic and convective element vertical motions are small scale phenomena, implying action on a given air parcel for only a short time whereas the large scale vertical motion field slowly displaces a given parcel over an extended period. Thus each may have a considerable influence on the total precipitation process.

Matveev[20f] has reported an interesting phenomenon associated with Cumulonimbus (Cb) cloud. He has reported from the data of 26 cases in the Kiev region in Ukrain that the precipitation exceeded the cloud precipitable water content on the average by a factor of 8.8 (with variations between 1.8 and 16.9). The water reserves of a Cb cloud are replenished every 7-12 minutes. The above data show that the amount of precipitation from the Cb cloud systems during their time of existence exceeds by about one order of magnitude of their water content at any given moment. This means that the water is replenished many times during the cloud's life period.

The above observation confirms the fact that the forecasting of the amount of rainfall associated with the convective cloud is a very difficult proposition if not impossible. As such QPF itself for any type of precipitation is a very difficult task.

Moreover, unlike other meteorological parameters rainfall is a highly variable quantity. Precipitation amounts are seldom representative and a few rain gauges do not constitute an adequate sample of a large area for quantitative purposes. Moreover, there are several methods for determining the average depth of precipitation and the amount computed by one method differs considerably from that computed by another method.

Acknowledgement

The author would like to thank the Director-General of Meteorology, Indian Meteorological Department, New Delhi, for kindly permitting him to publish this paper in this journal. Thanks are also due to M/s. James D'Souza, S M Jamadar, S G Nanaware, C M Shukla and P S Chougale for their help and cooperation.

References

1 A Belloq *Operational Models and Quantitative Precipitation Forecasts for Hydrological Purposes and Possibilities of an Intercomparison* World Meteorological Organisation Geneva Switzerland (1980)
2 K P Georgakakos and R L Bras *Water Resource Res* 20 (1984) No 11
3 K P Georgakakos and R L Bras *Water Resource Res* **20** (1989) No 11, 1597-1610
4 K P Georgakakos and M D Hudlow *Bull Am met Soc* **65** (1984) No 11, 1180-1200
5 D S Upadhyay, N Y Apte, S Kaur and S P Singh *Mausam* **37** (1986) No 3, 369-372
6 P N Sen *Mausam* **42** (1991) No 2, 201-204
7 J P Gerrity *The LFM Model—1976, A Documentation NOAA Tech Memo NWS NMC* (1977) 60

8 J E Newell and D G Deavan *The LFM-II Model* 1980 *NOAA Tech Memo NWS NMC* (1981) 66

9 K P Georgakakos *Stochastic Dynamic Short-term Prediction of Space-time rainfall* Presented at the Chapman Conference on Modelling of rainfall field, Caracas Venezuela March 24-27 (1986)

10 K P Georgakakos and P N Sen *A Stochastic-Dynamic rainall Prediction Model for use with surface and upper air Meteorological Data* (1986) (*unpublished*)

11 T H Lee and K P Georgakakos *J geophys Res* **95** (1990) No D3 2113-2126

12 P K Das *Attenuation of Vertical Velocity in the Atmosphere IMD Sci* (1969) Rep No 114

13 G J Haltiner and Martin F L *Dynamical and Physical Meteorology* McGraw Hill Book Co Inc New York (1957) 55

14 P R Lowe and J M Ficke *Tech Pap* No. 4-74 *Environ Prediction Res* (1974)

15 H R Pruppacher and J D Klett *Microphysics of Clouds and Precipitation* D Reidel Pub Company Dordrecht Holland (1980) 625

16 G K Sulakvelidze *Rainstorms and Hail* Translated from Russian by the Israel Program for Scientific Translation, Jerusalem (1969) p 19

17 J R Holton *An Introduction to Dynamic Meteorology* Third Edition Academic Press New York (1992) p 19

18 J O Rhea *Orographic Precipitation Model for Hydrometeorological Use, Ph.D Dissertation,* Colorado State Univ (1978) Atmos Sci Pap No 287

IMPORTANCE OF MOUNTAIN WAVES IN AVIATION AND WEATHER HAZARDS ASSOCIATED WITH IT

U S DE

India Meteorological Department, Shivajinagar, Pune-411 005, India

(*Received* 16 *August* 1993; *Accepted* 19 *August* 1993)

The paper summarises the importance of mountain waves and their impact on aviation with particular reference to the Indian scenario. The lee waves occur on different space scale. In the scale of waves ranging between 10 and 100 km, the flow is non-rotating and hydrostatic. Studies using steady state two dimensional models using perturbation technique are in general successful in explaining most of the salient features. However, nonlinear effects are to be considered for explaining the phenomena of hyraulic jump and upstream blocking, which are associated with strong down slop winds off steep mountain barriers. Work done in India and its relevance to aviation meteorology has also been discussed. The paper also summarises important factors which can be used for providing adequate guidance for safer air navigation in the air routes overflying orographic barriers.

Key Words: Mountain Waves; Hydraulic Jump; Rotor Clouds, Froude Number; Clear Air Turbulence

Introduction

Mountain waves belong to the category of atmospheric waves which interest both theoretical as well as operational meteorologists. Early aviators were largely aware of the hazards in flying over uneven terrain. From the view point of aeronautical meteorology, mountain waves account for some disastrous accidents in the past years. However, for the last four decades, a slow but steady collection of the information and observation together with adequate theoretical work have made it possible to overcome such hazards and make flying safer from this angle.

Theoretical Aspects

The atmosphere is capable of sustaining a large number of wave phenomena. We may classify them according to the nature of the displacement i.e., transverse or longitudinal or depending upon the nature of the restitutional forces i.e. gravity, rotation of the earth, inertia etc.

Mountain waves are essentially a class of gravity waves. However, the importance of the other terms like the compressibility, Coriolis force etc. are also to be considered for the appropriate scales of motion. The measure of significance of gravity is reflected if we consider the Froude Number

$$F=\frac{|U|}{\sqrt{gL}},$$

where U is the characteristic velocity, L is the characteristic length and 'g' is the acceleration due to gravity.

Historically, the origin of the problem can be traced back to the works of Rayleigh[1] and Kelvin[2]. They considered the flow of an incompressible inviscid and homogeneous fluid over an obstacle at the bottom. However, the corresponding problem in the earth's atmosphere is much more complex. This is due to thermal stratification of the atmosphere and the effect of the earths rotation.

Quenny[3] performed a sort of scale analysis of the atmospheric flow across a mountain barrier and showed that for an adiabatic perturbation the wave equation can be written as

$$\left(1-\frac{f^2}{K^2U^2}\right)\frac{\partial^2\psi}{\partial z^2}-\left(\frac{g}{C^2}+\beta\right)\frac{\partial\psi}{\partial z}+\left(\frac{g\beta}{U^2}-K^2\right)\psi=0 \qquad \ldots(1)$$

where

f is the coriolis force

ψ is the stream function

$$\beta=\frac{1}{\theta}\frac{d\theta}{dz}$$

θ is the potential temperature

C the speed of sound

and K the wave number.

He showed from physical consideration that there could be three different scales of mountain waves depending upon the value of and the other forcing terms:

if (*i*) $K>\sqrt{g\beta/U}$

(*ii*) $\frac{f}{U}<K<\sqrt{g\beta/U}$ $\qquad \ldots(2)$

and (*iii*) $K<\frac{f}{U}$.

In the first case motion is nonhydrostatic and non-geostrophic. While in the second case the motion is hydrostatic but non-geostrophic. In the last case, the motion is hydrostatic and quasi geostrophic.

In the atmosphere all the three types of mountain waves are possible, but the scales of waves from 1 to 100 km which fall under the category (*i*) and (*ii*) are of major concern to aviation. These waves are frequently observed and reported and can be approximated well by two dimensional linear models.

Most of the early investigation using the two dimensional linear models could explain usual features like the lee wave clouds and the wave motion on the leeward side of the mountains; Scorer[4], Doos[5], Fritz[6], De[7]. Organised field observations by Kuettner[8], Holmboe and Klieforth[9], Vergeiner[10] *et al.* and Küettner and Lilly[11] and recent experiment ALPEX have shown that the air motion in a mountainous terrain is highly modified and is often complex in nature. Broadly speaking two classes of flows have been documented—

(a) Wave motion.
(b) Hydraulic Jump type motion.

Wave Motion

Linear equations are adopted to accommodate a number of simplifying assumptions in the governing equations to obtain analytical solutions. Scorer[4], Doos[5], Palm and Foldvik[12] and Sarker[13] adopted this technique. Most of these studies used a two-dimensional adiabatic frictionless flow with the undisturbed flow being a function of the vertical co-ordinate only. By using the perturbation technique (where large amplitude and non linear effects are eliminated) they showed that a second order equation of the type

$$\frac{\partial^2 W}{\partial x^2}+\frac{\partial^2 W}{\partial z^2}+f(z)W=0 \qquad \ldots (3)$$

is obtained, where,

W is the perturbation vertical velocity and $f(z)$ is a complex function of stability and wind shear given by

$$f(z)=\frac{g(\gamma^*-\gamma)}{U^2 T}-\frac{1}{U}\frac{d^2U}{dz^2}+\left(\frac{\gamma^*-\gamma}{T}-\frac{g}{\varkappa RT}\right)\frac{1}{U}\frac{dU}{dz}-\frac{2}{\varkappa RT}\left(\frac{dU}{dz}\right)^2-\left(\frac{g-R\gamma}{2RT}\right)^2 .$$

where U, T are the undisturbed wind and temperature in the basic flow, γ and γ^* respectively lapse rate and the dry adiabatic lapse rate, R is the universal gas constant and $\varkappa = C_p/C_v$.

In the linear steady state models it is assumed that the perturbations are small and the product of perturbation quantities can be neglected.

Though the linear small amplitude theory is capable of explaining the "lee wave phenomenon" but important nonlinear effects like 'rotor' phenomenon and hydraulic jump type flow remain beyond its perview.

Non linear effects were introduced in the studies by Long[14,15]. He considered a stratified invicid but incompressible flow over a small barrier. Several other authors pursued studies following Long. Drazin and Moore[16], Miles[17,18], Miles and Huppert[19] and Davis[20] supported the initial results of Long. The phènomenon of upstream "blocking" and the "hydraulic jump" at very small Froude Number and for large barriers was confirmed.

Earlier Foldvik and Wurtele[21] considered the nonlinear effects by considering a time-dependent equation.

For the class of flow for an ideal fluid, the Froude number is used to discriminate the following up types of terrain induced motion.

(i) If $F>>1$ there are no lee waves,

(ii) If $F \rightarrow 0$ upstream blocking is complete

and for the intermediate range of values, there is partial blocking on the windward side and hydraulic jumps in the lee of the barrier. Since the Froude Number depends upon the mountain height also under typical atmospheric conditions these are observed when the mountain height exceeds 0.5 to 1 km.

Scorer[22] showed that the condition for occurrence of rotors to be

$$\frac{\partial W}{\partial z} > 1 \text{ or } < -1, \qquad \ldots (4)$$

which implies backward sloping streamlines. Rotors therefore tend to develop when the wave amplitude increases where the slope of the vertical profile of W is the largest.

Observational Aspects

As stated in the earlier section the effects of topography on air motion operate on a wide range of scales. There is as such an hierarchy of circulation systems associated with them. Fortuitious orientation of mountain barriers over the Indian subcontinent allows us to see all these effects in different space scales. Investigations by Das[23] refers to the planetary scale, while those by Sarker[13], De[7] and Sinha Ray[24] refer to the mesoscale. On the other hand investigations by Ghosh & Mukherjee[25] relate to the local circulations induced due to orography.

Early observations regarding mountain waves were based on visual observations of the leewave clouds and experience of pilots of sail planes and gliders. The Sierra wave project which was carried out over the Sierra Nevada Mountain ranges in 1951-52 was perhaps the most ambitious observational programme. Using radiosonde, time lapse photographs double theodolite balloon ascents and powered aircraft flights and sail planes observations a valuable data set was collected. The main results of the experiment showed that in case of powerful waves the amplitude could reach upto 2000 m at 6 km, height with vertical velocities ranging between +12.5 m to −9.5 m/sec. Often a single wave was reported in the stratosphere though the main wave activity was confined to the troposphere. During the last two decades more quantitative observations have been collected through the use of sophisticated ground base equipments in conjunction with instrumented aircraft observations. These have been summarized by Klemp and Lilly. A typical case of moderate lee wave (17 Feb. 1970) conditions as reported in their study relating to the Rocky mountains in Colorado is shown in Fig. 1. In another case of a more severe wave situation reported on 11 Jan 1972, the amplitude of the lee wave was of the order of 6,000 m and the associated vertical velocity of 20 mps was reported. The undisturbed flow conditions for this case showed presence of an inversion layer above the mountain top level and horizontal wind speed of 20 mps at the mountain top level. These facts highlight the importance of predicting such extreme situation and its impact on aviation.

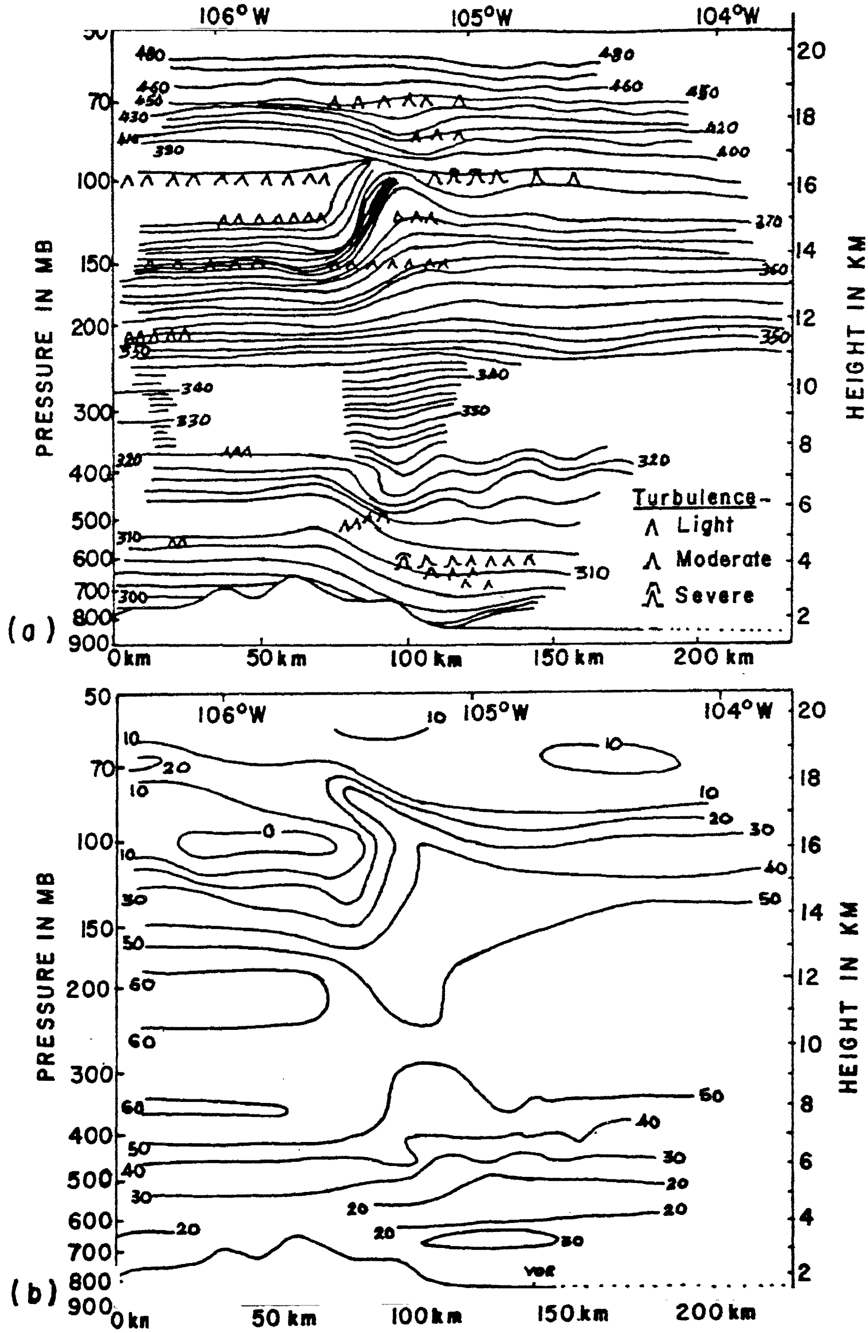


Fig 1 (*a*) Potential temperature cross section for 17 February 1970. Solid lines are isentropes (K)
(*b*) Westerly wind component cross section for 17 February 1970. Isotachs are in ms^{-1}

Theoretical and observational studies carried out in India have indicated that strong mountain waves occur under favourable condition during winter over the Assam-Burma hills, the Western and Central Himalayas and over the Western Ghats. The observed wavelength of the lee waves using satellite imagery for the Assam Burma hills ranged between 17-34 km. Lee waves for the Himalayas ranged between 10.2 to 22.3 km. in wavelength. Theoretical computations using linear two dimensional mountain wave model support these observations. For the Western Ghats wave length during winter ranged from 25.1 km to 78.5 km and during the southwest monsoon season between 19.2 to 31.7 km. Though the direct verification for the Western Ghats is not available yet aircraft reports on similar situations by and large support these figures.

The associated vertical velocities for the Assam Burma hills were of the order of 0.2 $\mathrm{m.sec^{-1}}$ to 4.3 $\mathrm{m\ sec^{-1}}$ during winter. Over the Western Ghats it ranged from 0.4 to 5.7 $\mathrm{m\ sec^{-1}}$. These observations clearly indicate the need for a closer study of the phenomenon over India. Major air routes which exist in these areas and any future expansion of air routes should take into account this problem and its impact on the routine operations.

Factors Influencing the Airflow Over the Orographic Barriers

Airflow over mountains in the horizontal scale of 1-100 km has a significant influence on the weather in the area. Aviators are thus most interested in these scales of motions which are directly induced by orography. In the text that follows an attempt has been made to highlight the salient features relating to this aspect of problem. The behaviour of the airflow in this spatial domain depends on the following factors:

(*i*) the vertical wind profile; (*ii*) the stability structure; and (*iii*) the shape and size of the obstacle.

From the point of view of aviation the vertical currents associated with the waves are of crucial importance. These waves can produce severe turbulence in which the aircraft is tossed up and down. On December 18, 1952, a Viking aircraft approaching the mountain barrier north of Madrid first fell by 2.7 km and subsequently rose by 4.3 km as it crossed the barrier. The sequence was repeat-

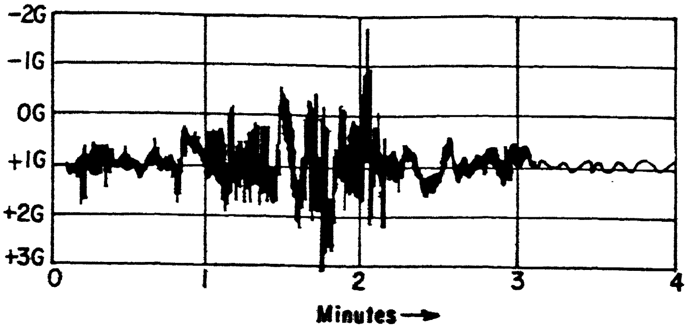


Fig 2 Flight recorder trace of vertical accelerations experienced during a severe turbulence encountered in clear-air turbulence (CAT) by a four-engine turbojet (B-747). This turbulence was generated by a severe mountain wave off the rocky mountains east of Helena, Montana

ed three times on the leeward side indicating that the aircraft was caught in powerful lee wave. In another case a four engine Turbojet (B-747) experienced severe turbulence due to mountain wave off the Rocky mountains east of Helena. The fluctuations in the vertical acceleration were more than ± 1g in a matter of one minute, as shown in Fig. 2. Over the Indian region the areas where mountain waves may be frequently encountered are Assam, the Western Ghats, J& K Hills and western Himalayas during winter and the South West Monsoon season.

De[26], Sinha Ray & De[27], Sinha Ray[24] and Tyagi and Madan[28] have extensively reported these through the use of satellite cloud imageries. Aircraft reports from the flights Madras-Bombay (Dec. 1964) and Colombo-Karachi (Dec. 1964) have reported turbulence associated with lee waves on the lee of the western ghats at roughly 8-10 km levels, Sarker[29]. A typical vertical velocity profile is shown in Fig. 3.

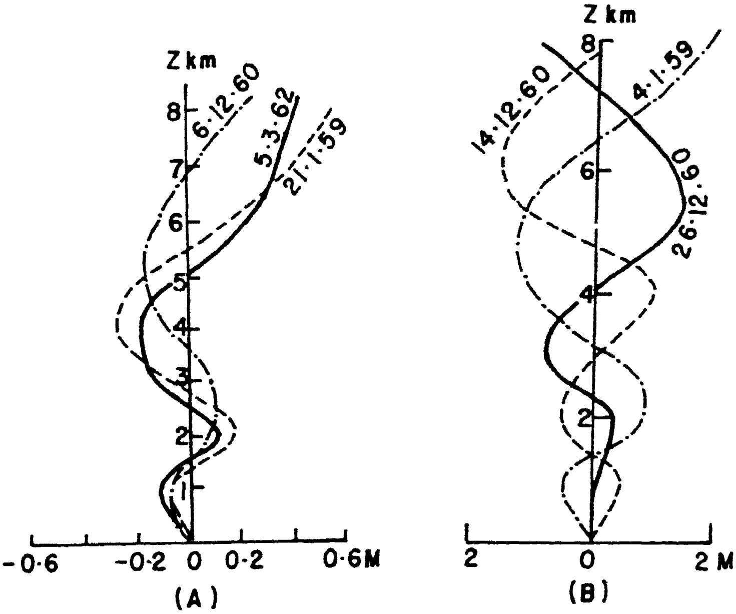


Fig 3 Variation of wave vertical velocity with height (Western Ghats)
(A) For short wavelengths and (B) For long wavelengths

Mukherjee[30] *et al.* have studied the occurrence of Clear Air Turbulence (CAT) reported by aircraft of the period of 1974-75. In this study they observed cases of CAT just east of Vishakhapatnam during winter months believed to be associated with orographic effects of eastern ghats.

The next important factor from the point of view of aviator is the amplitude of any lee wave. However, from an operational point it is difficult to predict it accurately. For a two-dimensional linear steady state model in x-z plane. Following Döös[5,31], the lee wave amplitude is given by

$$\zeta(x,z) = -\frac{U_0}{U_z}\sqrt{\frac{\rho_0}{\rho_z}}\,2\pi ab e^{-ak}\,\mathrm{Sin}\,kx\cdot\frac{J_m\left(2f_0^{1/2}\,e^{-\lambda z/2}\right)}{\frac{\mathrm{d}}{\mathrm{d}k}J_m\left(2\frac{f_0^{1/2}}{\lambda}\right)}, \qquad \ldots(5)$$

where $l^2 = f(z) = f_0\,e^{-\lambda z}$ ($f_0 < \lambda$ being constants) a is the half width and b the height of symmetrical bell shaped barrier and J is the Bessel function of the first kind.

Thus it depends on both the characteristics of the airstream as well as the features of the topography. In this respect, the half width of the mountain is a rather important factor. The dependence of the lee wave amplitude on the horizontal scale may be regarded as a "Resonance effect". If the power spectrum of the natural topography possesses a maxima near the natural wavelength of the air stream then the amplitude of the lee wave will be greater than otherwise. One can say if the horizontal scale of the barrier roughly coincides with the lee wavelength, the amplitude will be much larger than both, for broader or narrower mountains.

Descending currents on the leeward side of the mountain barrier close to the mountain top are a potential cause for accident if not adequately cared for. These descending currents create a cloud free zone due to adiabatic warming and an experienced pilot must avoid this pitfall. As mentioned in the earlier sections the occurrence of rotor zone due to closed circulation close to the ground (as a result of large tilt in the streamlines with height) is yet another source potential hazard to aircraft.

These should be avoided by aircraft descending or climbing near the leeward face of the mountains, when conditions for occurrence of lee waves are satisfied. These zones being close to the ground are often visible in the form of "Rotor Clouds" as shown schematically in Fig. 4. But if the air is not very moist such visible "alert" signs are missing thus making it more difficult to avoid them.

The occurrence of lee waves depends very crucially on the airstream characteristics. Use of Scorer's parameter l^2 to identify potential situation has been quite satisfactory provided the mountain top level wind ranges at least between 8-13 mps and it has a substantial component normal to the ridge line i.e., a southerly wind will be favourable for a East-west extending ridge and a westerly flow will be favourable for a north-south extending ridge. Scorer showed that no lee wave could occur if l^2 is constant with height. Introducing the concept of layered atmosphere he proved that for the occurrence of lee waves in a two layer model

$$l_1^2 - l_2^2 \geqslant \frac{\pi^2}{4h^2},$$

where l_1^2 and l_2^2 are the values of l^2 in the lower and upper layer respectively and h is the depth of the lower layer. In addition, an inversion layer or isother-

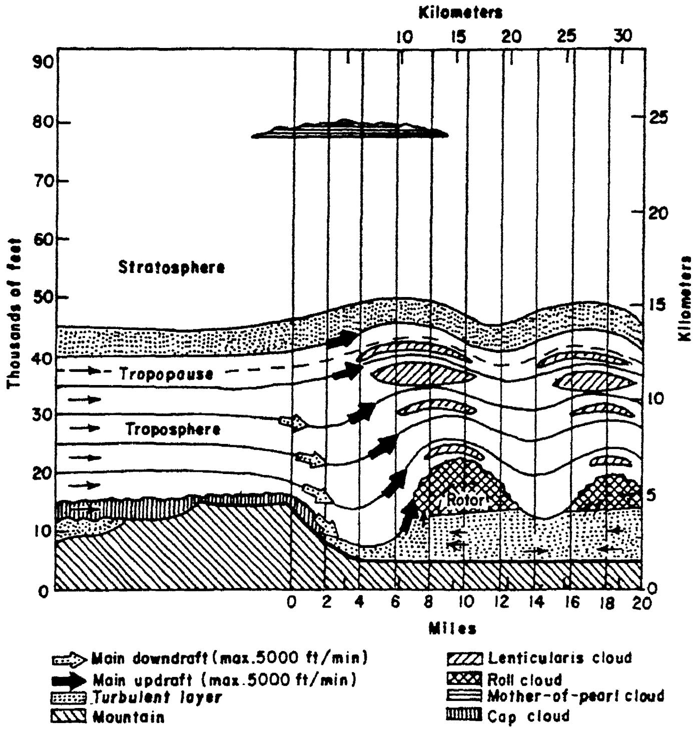


Fig 4 Idealized cross section of a mountain wave. Wind flow is left to right

mal layer in the lower or middle troposphere with a Jet stream (less stable) layer at 200-250 hPa are known to favour occurrence lee waves.

In a stratified atmosphere vertically propagating gravity waves occur when the intrinsic frequency is less than the Brunt Vaisala frequency. These waves propagate energy upward from the ground. On the other hand trapped lee waves occur down stream of the barrier when there is either (*a*) decrease in stability with height or (*b*) an increase in the wind speed with height or both.

Diurnal variation of wave length occur due to change of atmospheric stability between morning and evening. Changes in the synoptic pattern also affect the wave length. A rough estimation of lee wave length can be made by

$$L=2\pi\bar{U}\left\lfloor g(\gamma^{*}-\gamma)\right\rfloor^{1/2}$$

if the lapse rate is taken as 5°C/km

$$L=0.5\bar{U}.$$

The presence of an inversion layer is not a necessary condition for the occurrence of these waves; though the presence of an inversion layer produces a condition for occurrence maximum amplitude in a narrow zone of high static stability. Observations have indicated that the wave intensity decreases rapidly above the inversion layer. This fact is of considerable use for the forecasters to advise the pilots in choosing the flight level in a forecast—mountain wave situation.

Acknowledgements

The author is thankful to the Indian National Science Academy for offering this opportunity to contribute in their esteemed journal. He is also grateful to the Director General of Meteorology for his encouragement and consent to forward the paper. Dr P N Sen, Dr K C Sinha Ray, Shri D V Vaidya, Shri A R Murudkar and Shri P S Chougule, helped in various ways in preparing this review paper. The author wishes to acknowledge their help.

References

1 R J Rayleigh *Proc London math Soc* **15** (1883) 69-78
2 W T Kelvin *Phil Mag* **5** (1886) No 22, 353-357, 445-452, 517, 530
3 P Quenny *Misc Rep* No 23 Depts Met Univ Chicago (1947) 61 pp
4 R S Scorer *J R met Soc* **75** (1949) 41-56
5 B R Doos *Tellus* **13** (1961) No 3, 305-319
6 S Fritz *J appl Met* **6** (1965) No 4 31-37
7 U S De *Indian J Met Geophys* **22** (1971) No 3, 361-364
8 J Kuettner *Beitr Phys Frei Atmos* **25** (1939) 79-114
9 J Holmboe and H Klieforth *Investigation of Mountain Lee Waves and Air Flow over Sierra Nevada, Fin Rep Contract No Af* **19** (604) 728 Dept Meteorology Univ California, Los Angeles (1957) 209 pp
10 I Vergeiner *J R met Soc* **97** (1971) No 1 30-60
11 J Kuettner and D K Lilly *Weatherwise* **21** (1968) No 5, 180-197
12 E Palm and A Foldvik *Oslo Geof Publ* **21** (1960) No 6, 1-30
13 R P Sarker *IJMG* **16** (1965) No 4, 565-584
14 R R Long *Tellus* **5** (1953) No 1, 42-58
15 R R Long *Tellus* **7** (1955) No 3, 341-657
16 P G Drazin and D W Moore *J Fluid Mech* **28** (1967) 353-370
17 J W Miles *J Fluid Mech* **32** (1968) 549-567
18 J W Miles *J Fluid Mech* **48** (1971) 791-795
19 J W Miles and H E Huppert *J Fluid Mech* **35** (1969) 497-525
20 R E Davis *Fluid Mech* **36** (1969) 127-143
21 A Foldvik and M G Wurttle *Geophys J R astron Soc* **13** (1967) 167-185
22 R S Scorer *Symp Mountain Met Atmos Sci Pap* No 122, 75-101 Fort Collins Colorado State Univ (1967)
23 P K Das *IJMG* **15** (1964) No 4, 547-554
24 K C Sinha Ray *Mausam* **39** (1988) No 3, 257-262
25 S K Ghosh and A K Mukherjee *IJMG* **16** (1965) 429
26 U S De *IJMG* **21** (1970) No 4, 637-642
27 K C Sinha Ray and U S De *Mausam* **33** (1982) No 2, 265 p
28 A Tyagi and O P Madan *Mausam* **40** (1989) No 2, 181-184
29 R P Sarker *Forecasting Manual* Rep No IV-10 IMD (1974)
30 A K Mukherjee, C E J Daniel, A Thiruvengadathan and K Balchandran *IJMHG* **27** (1976) No 4, 377-384
31 B R Doos *Tellus* **14** (1962) No 3, 301-309

MESOSPHERE : COUPLING REGION BETWEEN THE UPPER ATMOSPHERE AND MIDDLE ATMOSPHERE

B H SUBBARAYA

Physical Research Laboratory, Ahmedabad 380 009, *India*

(*Received* 26 *August* 1993; *Accepted* 2 *September* 1993)

In recent years there has been increasing recognition of the role of the mesosphere as a coupling region between the upper and the middle atmospheres. A variety of dynamical perturbations generated in the lower regions propagate upwards and transport both energy and momentum as well as trace gases across the mesopause region. The altitude region 80km to about 100km is characterized by a high level of turbulence. The energy dissipated in this region is highly variable and this variability affects the structure and composition of the entire upper atmosphere significantly. Eventhough in recent years a great deal of observational data has become available for this region from MF and HF radars, Lidars and rocket data, most of it is from the middle and higher latitudes and there is very little data from the low latitude region which can be used to characterize the turbulence parameters of the low latitude mesosphere. The situation is expected to improve significantly in the coming years by the commissioning of the MST radar at Tirupati, Rayleigh Lidar at Ahmedabad and the Partial Reflection Radars at Thirunelveli and Kolhapur.

Key Words: Mesosphere; Coupling Region; Upper Atmosphere; Middle Atmosphere

Introduction

The mesosphere, region lying roughly between 60 and 90km altitude is the intermediate region between two distinctly different regions of the earth's environment:

(1) The upper atmosphere where absorption of solar energetic radiations of EUV and the X-rays by the major constituent O_2 (and to some extent N_2) dominates the atmospheric radiation budget, electric and magnetic fields play important roles in the physical processes taking place.

(2) The stratosphere—the altitude region of 20km to about 60km where absorption of solar ultraviolet radiation in the 2000A to 3000A region by the trace constituent in the atmosphere, namely ozone dominates the radiation budget and the physical processes are largely influenced by the dynamics forced from the weather related regions below.

This intermediate region has been and continues to be the least explored of all the regions in the earth's atmosphere. The altitude region is too high for balloon-borne and too low for satellite based *in-situ* measurements. It depends mostly on high altitude rockets and groundbased optical (airglow) and ratio probing techniques. Satellite-based techniques have been able to give some information about this region only in recent years. Further, there is very little direct absorption of solar radiation in this region and the energetics and dynamics (Table I) of this region is determined largely by transport of energy and

Table I

Energetics of the different regions of the atmosphere

Upper atmosphere/Thermosphere	$Z > 95$km
Absorption mostly by O_2 and to some extent by N_2	
Almost all the energy in the Solar spectrum short of $\lambda = 1750$A	30 ergs/sec/cm^2 Col.
Schumann Runge Bands 1750-1950A Lyman-alpha (1216A) and X-rays $\lambda < 20$A	100 ergs/sec/cm^2 Col.
Total	: 130 ergs/sec/cm^2
Stratosphere	: 20km $< z <$ 60km
Absorption by Ozone Solar UV (2000A-3000A)	1.8×10^4 ergs/sec/cm^2/col

The consequent atmospheric heating rates due to above are shown in Fig. 2.

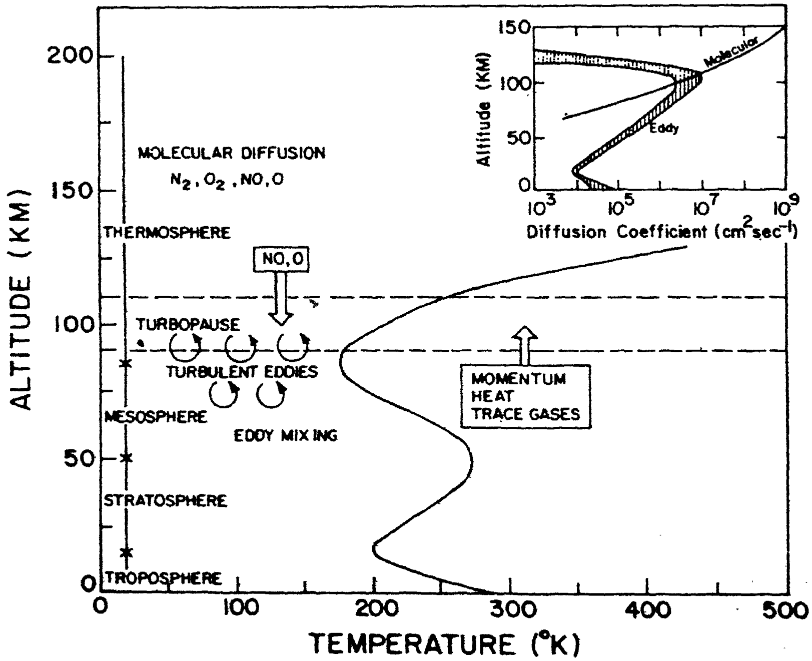


Fig 1 A schematic illustrating the differences in the characteristics of the middle and upper atmospheres

momentum from below. Because of the temperature minimum the mesopause was for a long time considered to be a buffer between the upper atmosphere and the lower regions without any interaction between the two. There are also other basic differences in the characteristics of the upper atmosphere from those of the regions below. Molecular diffusion coefficient increases with increasing altitude. Eddy diffusion coefficient is highly variable, both spatially and temporally, but all estimates show that it decreases rapidly above about 90km so that turbulence effects are negligible at heights above about 110km. Hence, the upper atmosphere is characterized by diffusive equilibrium with the structure and composition determined largely by the solar energetic radiation flux and the properties at the base of the region, while in the altitude region below about 90km the atmosphere is well mixed due to eddy mixing. The separation between the two occurs somewhere in the altitude region of 90-110km—the so called turbopause region (Fig. 1) where eddy viscosity roughly equals kinematic viscosity. Rocketborne vapour release and mass spectrometer measurements have shown that both the turbopause level and its properties are highly variable. Blamont[1,2] found that the turbopause level over Thumba is generally in the altitude range of 100 to 105km (Anandarao *et al.*, 1978)[3] found from a rocket experiment the turbopause at Thumba can sometimes be as low as 93km. Variability of turbopause level is linked with the turbulent energy input from below which depends on source strength and the propagation characteristics and hence could be highly variable. Fig. 3 (Subbaraya and Lal, 1983)[4] shows data from several rocket observations of mesospheric temperatures from Thumba which indicates a high degree of variability in mesospheric structure. Such a variability will affect the entire structure and composition of the thermosphere since the structure and composition of the base is an important input parameter in diffusion models of the thermosphere.

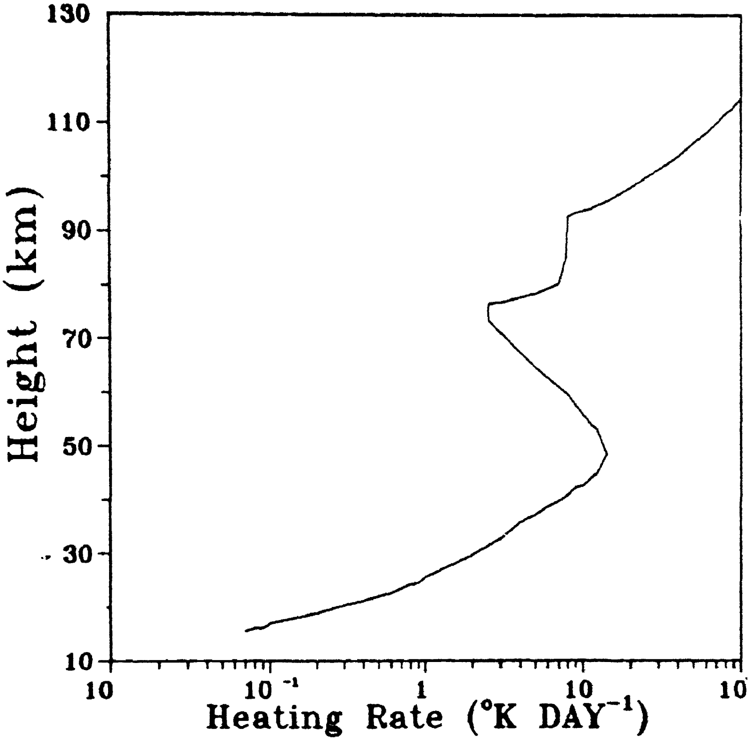


Fig 2 Heating rates in different regions of the atmosphere due to absorption of solar electromagnetic radiation

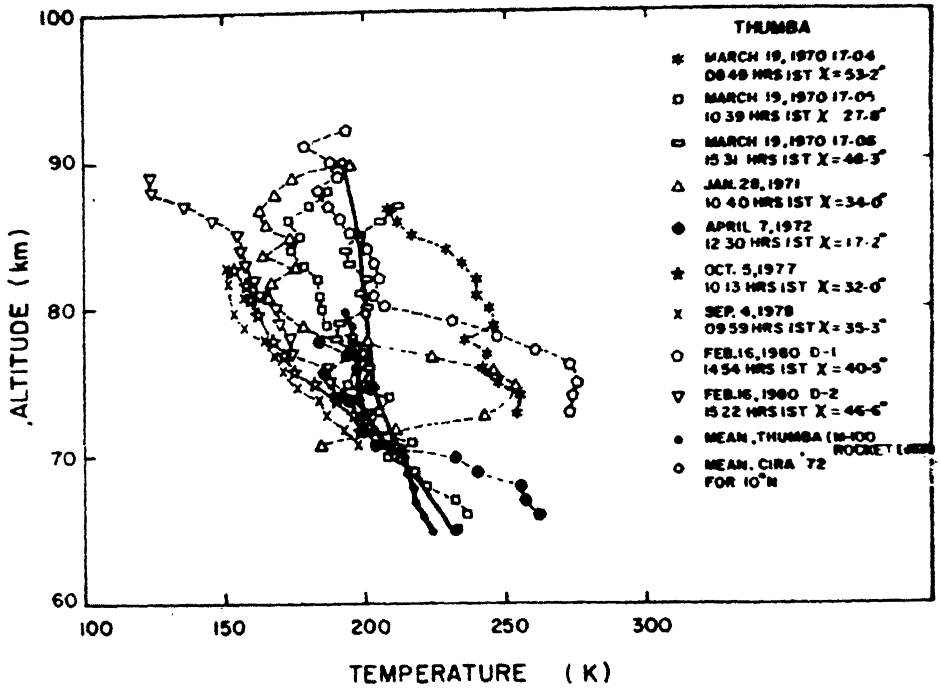


Fig 3 Mesospheric temperatures estimated from rocketborne Solar Lyman-alpha absorption measurements at Thumba[3]

Middle Atmosphere and Upper Atmosphere Coupling

Several developments in recent decades have forced the recognition of the mesosphere as an important region for study. As seen from Table I, the direct absorption of solar energy in the upper atmosphere is less than one percent of that in the stratosphere. So even a small leakage of energy from below, a leakage of the order of a fraction of a percent can cause significant perturbations in the energetics of the upper atmosphere. There has been plenty of evidence in recent years to support significant coupling between the middle and the upper atmospheres. Several studies show good correlations between changes in the ionospheric behaviour with changes taking place at stratospheric altitudes. A number of studies have shown that ionospheric perturbations, such as travelling planetary waves in the ionosphere are correlated with wave activity in the stratosphere[5-9,9a]. Brown and Williams[10] showed a correlation of E-region electron densities with 10mb pressures, especially in winter months. More recently, Luo Yi and Pei Rem Chen[11] have reported oscillations with periods of 2-2.5 days to 10-16 days in the F-region equatorial ionization anomally that are correlated with mesospheric dynamics. Further, some of these oscillations are shown to be modulated by QBO indicating a coupling all the way between the troposphere and the F-region ionosphere. Standard atmosphere models have been steadily lowering the base of the upper atmosphere from 120km to 90km from the first CIRA-61 to the models of the seventies and the eighties indicating the impact of middle atmospheric changes on thermospheric properties.

Coupling Processes

Coupling can take place between the middle atmosphere and upper atmosphere in several ways. The coupling process can be broadly classified under three headings:

1. Transport of energy and momentum by waves of different scale sizes and periods—Advective and Diffusive transport. The latter being due to both molecular diffusion and turbulence (eddy diffusion).
2. Transport of Trace Constituents—advection and diffusion, both upward diffusion from the stratosphere to higher levels, e.g., O_3, H_2O, CO and downward diffusion from the thermosphere, e.g., Nitric oxide and atomic oxygen depending on the source region.
3. Electrodynamic coupling—exists across all regions.

All the three processes are quite important. But the effects of trace gas transport are indirect, through the role the trace gases play in absorption of radiation and other photochemical interactions. Electrodynamic coupling will impact directly on the ionization parameters only. The most direct effects are through energy and momentum from the middle to the upper atmosphere and the most important is the transport across the mesopause region, by waves and turbulence. The following sections will discuss this in greater detail.

Waves and Turbulence in the Middle Atmosphere

Existence of waves in the atmosphere over a large spectrum of scale sizes from planetary scales (several thousands of kms) down to values of the order of a meter or less and periods from years down to fraction of a minute has been well documented from a variety of observations. Waves play an important role in the dynamics of the atmosphere and manifest themselves as fluctuations in mean winds, temperature, density, ionization as well as airglow intensities. The height region of 80-100km is the region where the effects are most pronounced. This is the seat of interaction between the upper and the middle atmosphere.

Wind measurements in the mesosphere made at different locations have been spectrum analysed by Vincent[12] to study how wind energy is distributed as a function of frequency. The result illustrated in Fig. 4 shows that waves excist in three distinct regimes, viz:

1. Long period waves with periods greater than 24 hours (transient planetary waves). They occur in three distinct period ranges: 10-20 days period with amplitudes 5-10m/sec and wavelength greater than 100km, 4-7 days period with amplitudes 5-30m/sec and wavelengths in the range of 25km to about 100km and roughly 2 day periods with amplitude 10-50m/sec and wavelengths 50-100km.
2. Medium period waves—12-24 hrs—most important are the semi-diurnal and the diurnal tides.

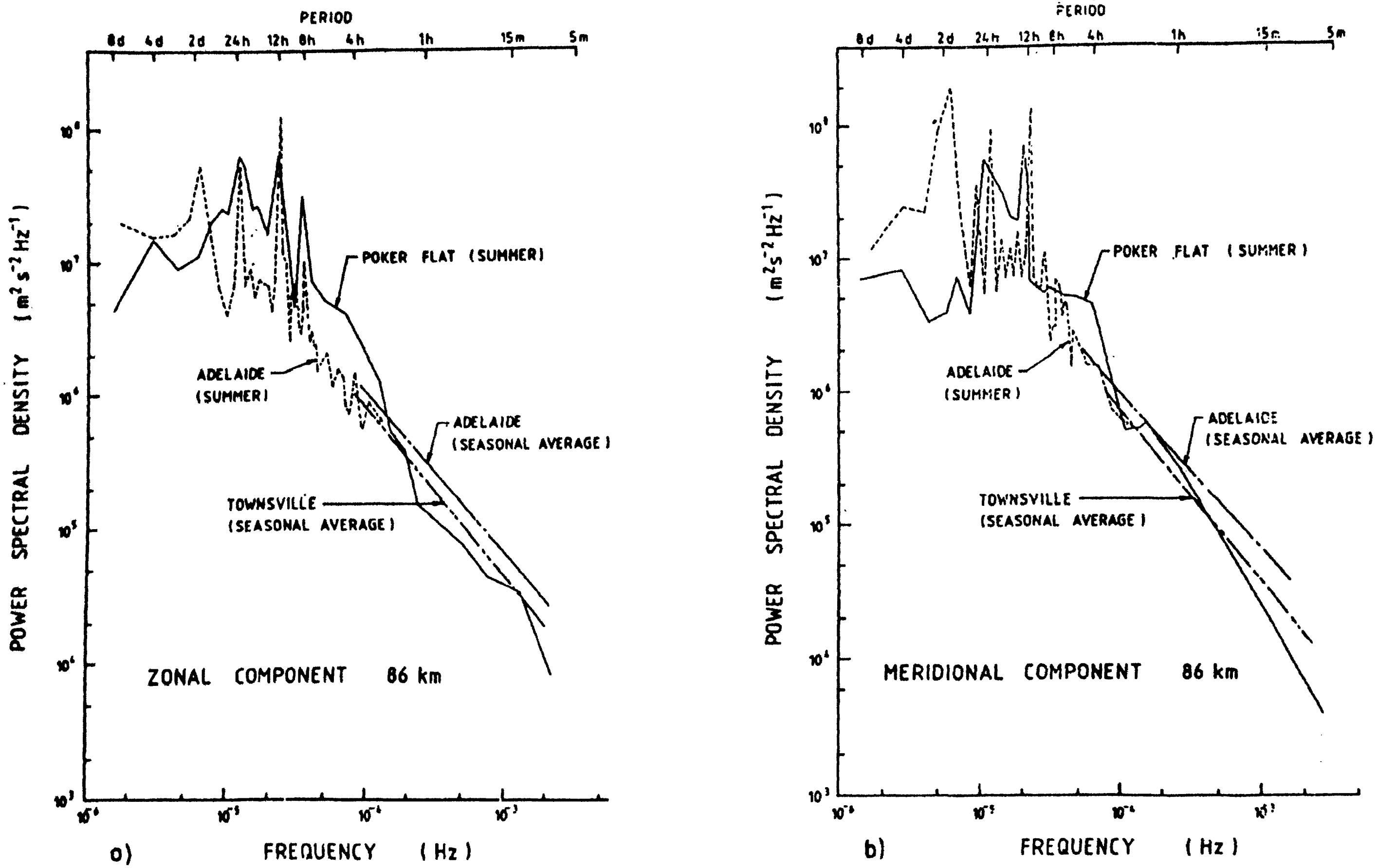


Fig 4 Spectral distribution of wave energy in the mesosphere obtained from a compilation of several VHF and MF radio propagation measurements (Vincent 1985)[12]

3. Short period waves periods less than 12 hrs. These are internal gravity waves. Spectrum extends down to a few hours with amplitudes generally decreasing monotonically with decreasing period. Spectral index is of the form f^{-n} with n values lying in the range 1.5 to 2.

Rocket measurements have been used extensively to study long period waves. There is plenty of evidence for the existence of Kelvin and Mixed Rossby waves. Raghavarao *et al.*[13a] have found waves with 30-50 days period at low latitudes. In recent years, there has been a good deal of work on the 2 day waves. Observational data[13,14] shows that these short period waves carry a significant fraction of the energy and momentum flux and play an important role in coupling the middle atmosphere to the upper atmosphere eventhough it is the long period waves that contain the largest energy densities. For the purpose of the present study, the most important wave motions are the diurnal and semi-diurnal tides and internal gravity waves eventhough in recent times waves of shorter periods have also been reported. The tidal modes (both diurnal and semi-diurnal) are generated in the troposphere and the stratosphere due to absorption of solar radiation by ozone and water vapour and propagate upwards. Internal gravity waves have periods shorter than those of the tides, but are similar in most other characteristics. Anandarao and Raghavarao[15] find that the intrnal gravity waves are most prominent in the 90-110km region They are generally generated in the lower regions. Once generated, they propagate upwards[16] and as they propagate they grow in amplitude (roughly as $\rho^{-1/2}$) in the absence of damping. As the amplitude increases, at some altitude the wave becomes unstable and gives rise to turbulence. Turbulence in the atmosphere could also be due to shear instabilities, convective overturning and critical level interaction when several waves are simultaneously present[17].

Internal gravity waves have been most extensively studied using MST radar data. At a given location the wave amplitudes are generally the same both for zonal and meridional components indicating that the perturbation is on the average isotropic. The wave characteristics vary with latitude and season. Greatest wave activity is in winter high latitudes (amplitude 30°K). Mid-latitude values are of the order of 10°K. Vertical wavelength of gravity waves in the mesosphere is of the order of a few km. Rocket measurements of temperature, molecular oxygen and ozone distributions measured at Thumba (Figs 5, 6 and 7) illustrate some of the features. Horizontal wavelengths can be obtained only from airglow observations. Moreels and Herse[18] have made groundbased image intensifier camera observations of OH airglow intensities (in the 85-90km region) and report that horizontal wavelengths are of the order of a few hundreds to a thousand kilometers. Manson[19] estimates similar values for the horizontal scale sizes using vertical wavelength measurements from radar data and using the dispersion relation for gravity waves. Rayleigh Lidar observations at Haute Province in France[20], Lidar Sodium layer observations over Urbana[21] have also been used to study internal gravity wave characteristics at mesospheric altitudes. Vertical wavelengths are found to lie in the range 2km to about 15km and periods are of the order of a few tens of minutes to hundreds of minutes. Groundbased OH and Sodium night airglow observations at Mt.

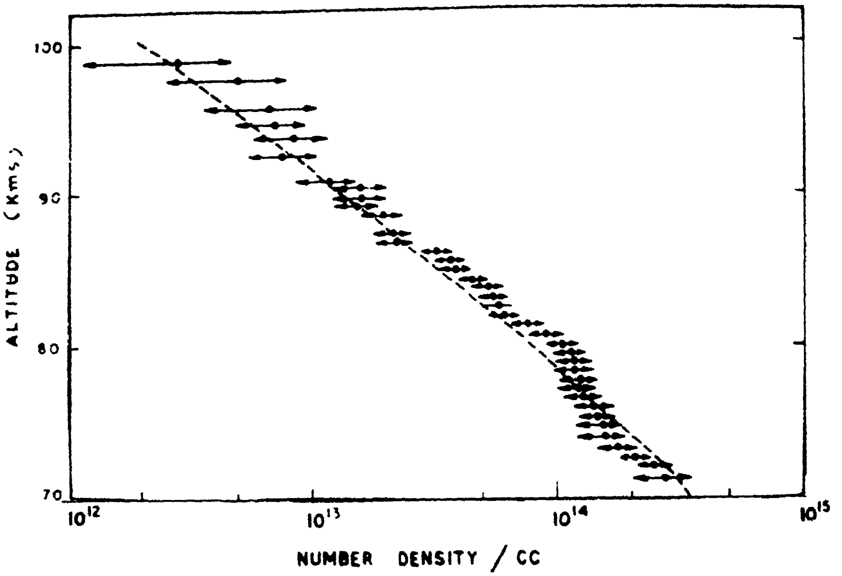


Fig 5 Wave-like perturbations in the vertical distribution of molecular oxygen concentrations measured at Thumba[4]

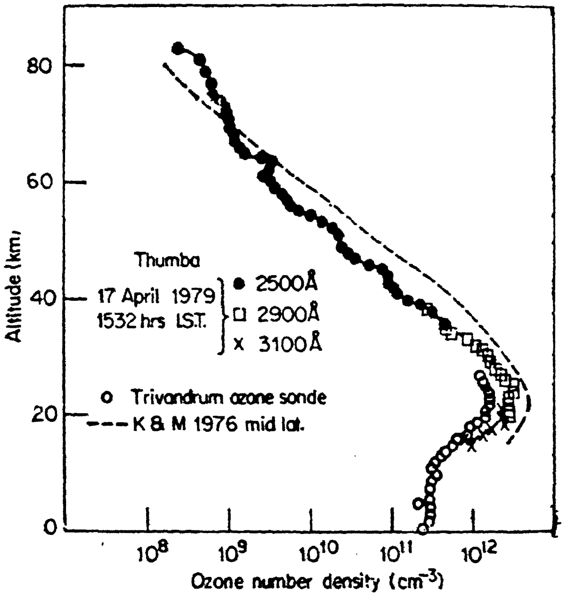


Fig 6 Wave-like perturbations in the vertical distribution of ozone concentrations measured at Thumba[4]

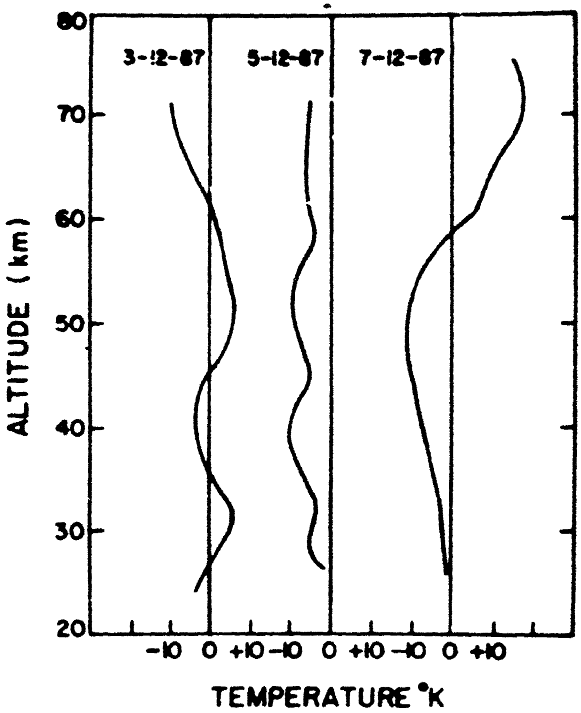


Fig 7 Wave-like perturbations in the temperature structure over Thumba measured during Dec. 1990 rocket campaign[4]

Abu[21] have revealed frequent wavelike features with periods of the order of a few hours. Nakamura *et al.*[22] report that the mesospheric gravity waves exhibit similar features both in the northern and southern hemispheres from a comparative study of the Japanese MU radar (35°N) data with that of the Adelaide MF radar (35°S). This suggests that eddy dissipation processes due to turbulence at both the locations are similar. This is an intriguing suggestion considering that several studies[23,24] have shown considerable asymmetry in the characteristics of the tides in the northern and southern hemispheres.

Vertically propagating gravity waves including some of the tidal modes get trapped in the mesospheric region and eventually the energy gets dissipated. This can cause major perturbations in the mesospheric and lower thermospheric energy budget. Hines[25,26] suggested that the heating rate due to gravity wave dissipation could be as large as 10°K/day in the 90-95km region. This can be compared to the value of 1-2°K/day due to direct absorption of solar radiation (*See* Fig. 1). Vertically propagating gravity waves play an important role in the dynamics of the intermediate region. In fact, the meridional temperature distribution and even the mean wind profile is affected by turbulent diffusion[22].

Zimmermann and Rosenberg[28] studied wind and wind shear data from about 70 rocket measurements in the 90-150km region. Assuming a simple model of freely propagating wave dissipated by viscous energy continuously from below they found that the data is in agreement with kinematic viscosity considerations at altitudes above about 110km. Below this altitude the data would lead to very large (impossible) values of viscosity. Hence, the altitude region of 90-110km is suggested to be dominated by turbulence. Detailed spectral analysis of the waves at different altitudes enabled them to study the relative importance of eddy viscosity and molecular viscosity which confirmed that eddy diffusion dominates the 90-110km region whereas above 110km molecular diffusion takes over. These results are in qualitative agreement with Fig. 2. The energy deposition estimated from their wave data in the 90-110km region is $1\text{-}2\times 10^{-6}$ ergs/cm^2/sec which is much larger than the Schumann Runge absorption in that region.

Vincent (*loc. cit.*)[13] and Meek and Manson have extensively studied the Partial reflection radar data at Adelaide and Saskatoon respectively. Significant vertical motions (of the order of a few meters per second) have been found in the 75km to 100km region. Short period waves ($\tau <$ 1hr) with vertical wavelengths 10-12km are frequently found. The data have been used to determine eddy diffusion coefficients and energy flux which is of the order of 10^{-2} W/km^2. Lindzen and Blake[27a] estimate that the most important tidal mode—the fast propagating diurnal tide would breakdown ($R_0 < 0.25$) at altitudes of 80 to 90km and generate turbulence upto an altitude of 105-110km. In a crude model of turbulence in the mesosphere in which eddy diffusion cancels the exponential growth rate it is possible to estimate the eddy diffusion coefficient which gives Deddy values of the order of 10^6 cm^2/sec. in the mesosphere. This particular mode is especially important at the equator since the upward flux of energy carried by this wave according to classical theory is 6-8 ergs/cm^2/sec at

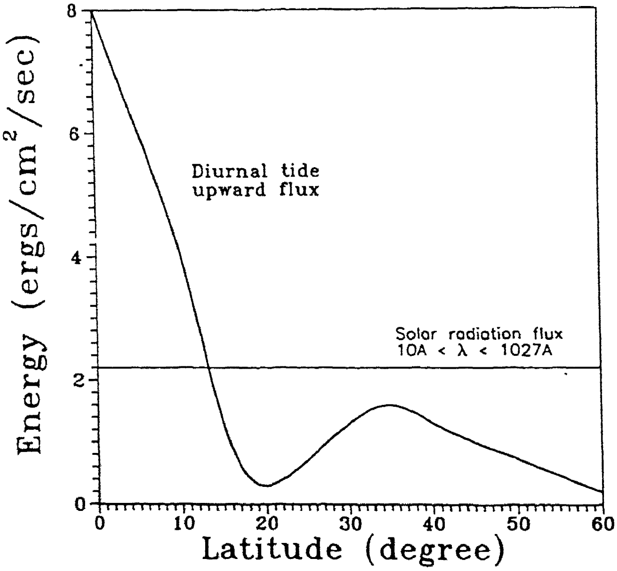


Fig 8 Upward flux of energy due to solar semidiurnal tide compared with the energy input du to solar EUV absorption[27]

the equator which is much larger than the solar radiation flux available in the EUV and X-ray region which is ~2 ergs/cm^2/sec. Hence, the equatorial region is most important for study of upward flux due to tidal and gravity wave forcing. Eventhough there has been significant improvements in tidal theory since these early estimates[29,30], they do not affect the above conclusions.

Turbulence in the Mesosphere through Ionization Irregularity Measurements

At altitudes below about 80km where the atmosphere is collision-dominated, electron and ion density irregularities can also be used to study neutral atmospheric turbulence. VHF forward scatter[31] and backscatter radar measurements[32] were amongst the first techniques used to study mesospheric turbulence. These pioneering studies showed that there are turbulent ionized layers in the mesosphere with scales of 100-200 meters and electron density fluctuations $(\Delta N^2/N^2) \sim 10^{-4}$. Partial reflection radars have also been used to study the ionization irregularities at D-region altitudes[33]. However, major advances in our understanding of the ionization irregularities became possible when *in-situ* high frequency response probes were used on rockets. First such measurements were made by Prakash *et al.*[34,35] at Thumba. They showed that in the equatorial region fluctuations in ambient electron densities exist at all heights in the altitude region of 60km to about 160km most of the time but

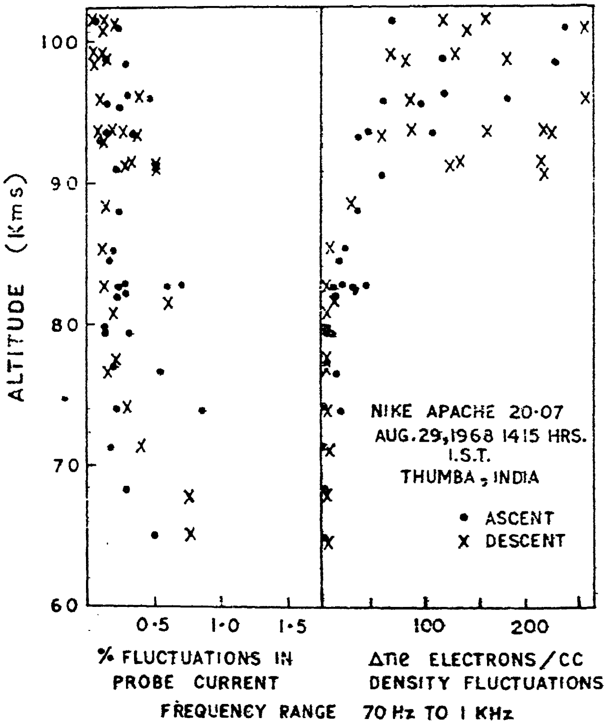


Fig 9 Early rocket measurements of ionization irregularities over Thumba (Prakash *et. al.*, 1971)[35]

are dominant in the altitude region of 80km to about 110km. Fig. 9 shows the results of early measurements. In the equatorial regions at altitudes above about 90km the irregularities present are mostly due to plasma processes[36] where electrodynamic parameters are important and hence are not very helpful for study of neutral atmospheric turbulence. But at altitudes below about 80km electric and magnetic field effects are small and the irregularities observed, which had a spectral index of $-5/3$ are attributed to neutral atmospheric turbulence.

Measurements similar to those of Prakash *et al.* were later conducted by several other workers[37,38]. There have even been attempts to make near simultaneous rocket and radar measurements[39]. Fig.. 10 illustrates data from two near simultaneous rocket probe and MST radar measurements with the rocket trajectory actually passing through the MST radar beam path. Enhanced backscatter signals from the radar collocated in altitudes with region of electron density irregularities as seen by the rocket probe are evident from the figure. Stoltzfus *et al.*[38] used Langmuir probe data from nine rocket flights to study turbulence parameters in the 60 to 90km region. Irregularities with measurable amplitudes were found in five of the nine flights. They were subjected to power spectral analysis. Examination of the maximum wave number for which the power law provides a good fit suggests that the cut off wave number decreases monotonically with altutide (Fig. 11). Comparison of rocket and Jicamarca VHF radar data shows that the radar operates in the viscous range of turbulence atleast in the upper part of the mesosphere. They further found that the energy dissipation rates calculated from both the rocket and radar data are in good agreement.

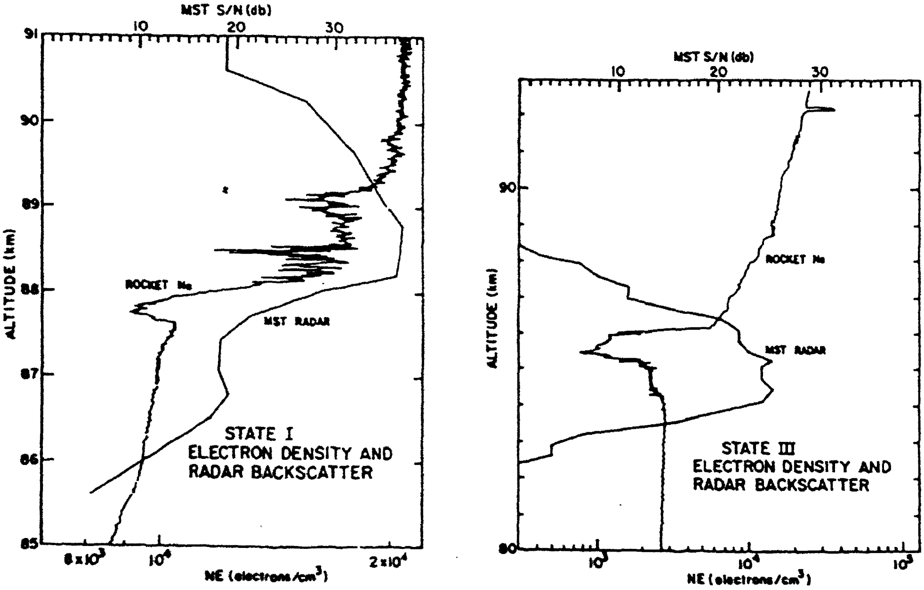


Fig 10 Simultaneous rocket probe and MST radar measurements of mesospheric irregularities (Philbrick *et al.*, 1984)[39]

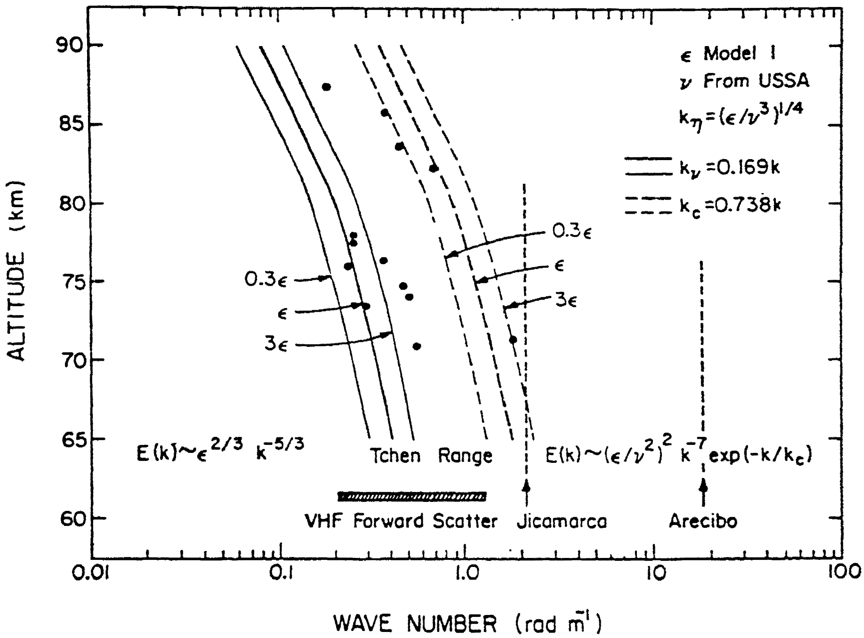


Fig 11 Altitude variation of the cut off wave number beyond which the $-5/3$ power law is applicable. Figure shows that the VHF backscatter radars at Jicamarca and Arecibo operate in the viscous region

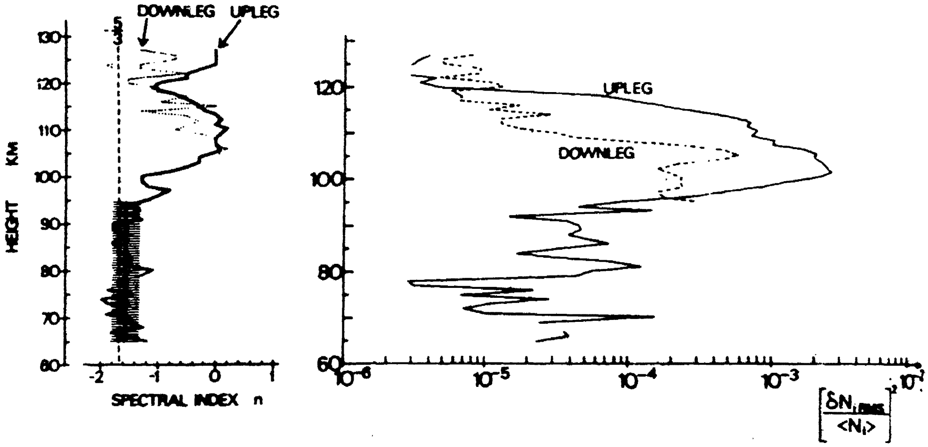


Fig 12 Rocket measurement of ion density irregularities over Andoya. $-5/3$ power law holds upto an altitude of 95km

Thrane and Grandl[39] found irregularities with spectral index $-5/3$ attributable to neutral turbulence at altitudes upto about 95km over Andoya (Fig. 12). Chakrabarty *et al.*[40] used a similar technique at Thumba and found that the ion density fluctuations in the altitude range 65 to 97km followed a spectral index close to $-5/3$. They estimated eddy diffusion coefficient values to lie in the range 10 to 100m^2/sec. in the altitude range 70 to 80km. They further used these values to determine the energy dissipation rates at these altitudes and obtained values in the range 0.01 to 0.1 W/kg. Sinha[41] examined the electron density irregularity data from a number of rocket experiments conducted earlier at Thumba to study the characteristics of neutral turbulence at low latitudes. Irregularities that can be attributed to neutral turbulence were found in the altitude region 65 to 80km. Estimates of energy deposition by turbulence in this altitude range vary from 0.07mw/kg to 0.62mw/kg which are comparable to the model values of Johnson[42], but are somewhat lower than those of Chakravarty *et al.*[40]

Concluding Remarks

More than half of all available data in the mesosphere is indicative of turbulence. In recent time Rayleigh Lidars capable of giving atmospheric densities and temperatures upto an altitude of about 65km and sodium resonance Lidars which can make measurements in the 70km to about 100km, altitude region, VHF backscatter, partial reflection and MST radars have been deployed at several locations and the data from these has given rise to a quantum jump in our knowledge of the turbulence in this region. Turbulence parameters, scale sizes, intensity of turbulence as indicated by the eddy diffusion coefficient and the estimated energy dissipation rates exhibit a large range of variability both with latitude and season. Turbulence/Eddy deposition of energy in the mesosphere is dependent on energy input from below and can be highly variable. Fig. 13 from Hocking[43] illustrates the variability in the eddy diffusion coefficient and the energy dissipation rate from a compilation of the data from various sources. Such a variability affects the characteristics of the turbopause region and hence can have a significant impact on thermospheric structure and composition.

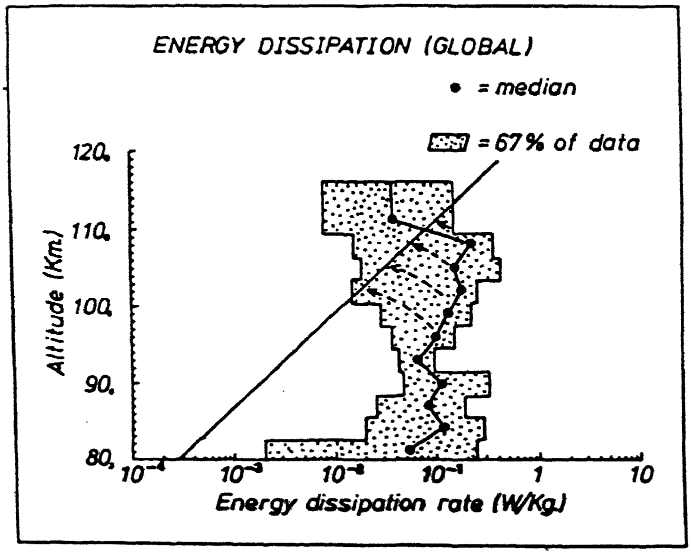


Fig 13 Variability in the energy dissipation rates due to turbulence at mesospheric altitudes (Hocking, 1985)[43]

Much of the observational data comes from middle and high latitudes. The sparse data available for low latitudes comes from sporadic rocket experiments. However, recently the National MST Radar Facility has become operational at Gadanki near Tirupati (13.5°N). Further, a Rayleigh Lidar is being set up at PRL, Ahmedabad (23°N). Two Partial Reflection Radars are being commissioned by IIG at Thirunelveli close to the dip equator and at Kolhapur (16.7°N). With these groundbased techniques, a systematic data base for the low latitude will become available and it should be possible to quantify mesospheric turbulence parameters over the low latitude Indian region in a more realistic manner.

Acknowledgements

Dr A Jayaraman was helpful in discussions regarding some parts of the manuscript. Acknowledgements are also due to Mr M N P Nair for secretarial assistance.

References

1 J E Blamont *Planet Space Sci* **10** (1963) 89
2 J E Blamont and J Barat *Ann Geophys* **24** (1968) 375
3 B G Anandarao, R Raghavarao, J N Desai and G Haerendel *J atmos Terr Phys* **40** (1978) 157
4 B H Subbaraya and S Lal *J atmos terr Phys* **45** (1983) 41
5 W G Beynon and G M Brown *Nature* **167** (1951) 1012
6 E A Lauter *Space Res* 7 (1967) 212
7 K Davis and J E Jones *J atmos Sci* **28** (1971) 254
8 G Chimonas and W R Pettier *J atmos terr Phys* **35** (1974) 821
9 V R Rao and B R Rao *Indian J Radio Space Res* **6** (1977) 136
9a D G Krishnaraju, M Sriramarao, B Madusudan Rao, C Purnachandra Rao, C Jogulu and B Ramanadan *J geophys Res* **86** (1981) 5873
10 G M Brown and D C Williams *J atmos terr Phys* **33** (1971) 1321
11 Yi Luo and Pei Rem Chen *J atmos terr Phys* **55** (1993) 1317
12 R A Vincent In: *MAP Handbook Vol* **16** (1985) 269
13 R A Vincent *J atmos terr Phys* **46** (1984) 119
13a R Raghavarao, R Suhasini, R Sridharan, B V Krishnamurthy and O P Nagpal *Proc Indian Acad Sci* **99** (1990) 413
14 D C Fritts *Res Geophys Space Res* **22** (1984) 275
15 B G Anandarao and R Raghavarao *Space Res Vol* **17** (1979) 263
16 C O Hines *Canad J Phys* **38** (1960) 1441
17 W J Hocking *Adv Space Res Vol* **17** (1987) 71
18 G Moreels and M Herse *Planet Space Sci* **25** (1977) 265
19 A H Manson *J atmos Sci* **47** (1990) 2765
20 M L Chanin and A Hauchercone *J geophys Res* **86** (1981) 9715
21 C S Gardner and D G Voelz *J geophys Res* **92** (1987) 4673
21a V R Rao and P V Kulkarni *Indian J Radio Space Res* **1** (1972) 181
22 T Nakamura, T Tsuda, S Fukao and S Kato *Geophys Res Lett* **20** (1993) 86
23 T Aso and R A Vincent *J atmos terr Phys* **44** (1980) 267
24 R A Vincent, T Tsuda and S Kato *J atmos terr Phys* **51** (1989) 609
25 C O Hines *Q J R met Soc* **89** (1963) 1
26 C O Hines *J geophys Res* **70** (1965) 177
27 R S Lindzen *J geophys Res* **86** (1981) 9707
27*a* R S Lindzen and D Blake *Geophys Fluid Dyn* **2** (1970) 31

28 S P Zimmermann and N W Rosenberg *Pap at COSPAR Sym* (1974) 134
29 H Titenbaum, C A Kelder and van Duin (1986) *J atmos terr Phys* **49** (1986) 413
30 F Vial and J M Forbes *J atmos terr Phys* **51** (1989) 663
31 R M Gallet *Proc IRE* **43** (1953) 1240
32 D K Bailey, R Bateman and R C Kerby *Proc IRE* **43** (1953) 1461
33 A H Manson, M W J Merry and R A Vincent *Radio Sci* **4** (1964) 955
34 S Prakash, S P Gupta and B H Subbaraya *Radio Sci* **4** (1969) 791
35 S Prakash, B H Subbaraya and S P Gupta **33** (1971) 129
36 S Prakash, S P Gupta, B H Subbaraya and R Pandey In. *Low Latitude Aeronomical Process, COSPAR Sym Sr* **8** (1978) 3
37 E V Thrane and R Grandl *J atmos terr Phys* **43** (1981) 179
38 R B Stoltzfus, S A Bowhill and L G Smith *Adv Space Res* **4** (1984) No 6, 143
39 C R Philbrick, D P Sipler, B B Balsley and J C Ulwick *Adv Space Res* **4** (1984) No. 6, 129
40 D K Chakrabarty, G Beig, S R Das and J S Sidhu *J atmos terr Phys*, **51** (1989) 19
41 H S S Sinha *J atmos terr Phys* **54** (1992) 419
42 F S Johnson *J atmos Sci* **32** (1970) 1658
43 W J Hocking In: *MAP Handbk Vol* **16** (1985) 290

Proc. Indian natn. Sci. Acad., **60**, A, No. 1, 1994, pp. 243-262.
© Printed in India.

TROPICAL IONOSPHERE AND RADIO COMMUNICATIONS

D R LAKSHMI
National Physical Laboratory, New Delhi-110012, *India*

(*Received* 3 *October* 1993; *Accepted* 26 *October* 1993)

The partially ionized layers of our upper atmosphere have facilitated traditional long distance communications for several decades because of their ability to reflect HF radio waves and take them to target areas thousands of kilometers away. The tropical ionosphere, protected from solar particle invasions, by a favourable geometry of the terrestrial magnetic field, offers many natural advantages. While these advantages encourage the continuance of HF communications in this region even in this present era of other reliable but expensive systems, it is becoming increasingly necessary to model and predict such debilitating factors as spatial and temporal gradients, day-to-day variations, storm departures and other short and long term trends to take full advantage of this rich natural resource.

With satellite based radio systems coming into vogue, the ionosphere assumed a different role. Frequencies above 10 GHz are plagued by tropospheric problems such as multipath fading, ducting and attenuation by rain and water vapour. But the lower frequencies are subjected to ionospheric phenomena, mainly scintillations and group delay. It is necessary to understand, quantify and predict these debilitating factors to aid in the design of robust radio systems that are either impervious or at least can reconcile with such phenomena.

The ionosphere also provides us with a vast and unbounded plasma laboratory with naturally occurring tracers such as magnetic storms or irregularities that can be used to study fundamental plasma phenomena.

Key Words: Tropical Ionosphere; HF Communications; Ionospheric Models; Scintillations

Introduction

The presence of ionized layers in our upper atmosphere is being exploited for inexpensive long distance communications and broadcasting for more than six decades now. It has also spawned a series of excellent research efforts to study the physics of weak plasmas and to understand the overall personality of our geo-environment. While we try to stretch this gift of a unique natural resource to the limit for ionosphere-aided communications, the realisation came more recently that this very ionosphere could be a natural hindrance to more sophisticated satellite-based systems.

The recent two or three decades have witnessed rapid advancements in the radio systems hardware, necessitating a better knowledge of the restrictions imposed by the propagation media on their performance. It is the media which essentially set the ultimate limits on the reliability of the communication. We will be limiting our discussion here mainly to tropical (equatorial and low latitude) ionospheric features as our interest is mainly on radio communications in this region.

The ionosphere is the ionized region in our upper atmosphere extending from about 50km to 1000km where sufficiently large number of free electrons exist to affect the propagation of radio waves. The ionosphere is produced essentially by the solar EUV and X-ray radiations and to some extent by energetic solar particles at high latitudes. Ionosphere can be divided into three distinct layers (regions) namely *D*, *E* and *F* and relative to background ionization in these regions, there can be patches of irregular structures embedded in them. (Fig. 1). Some of the irregularities are associated with phenomena like sporadic *E* (E_s) and spread *F* which can be identified from ionograms.

While both *E* and *F* layers support HF radio wave propagation over long distances the role of *D*-layer is mostly negative because of the large attenuation the HF radio waves undergo in this layer during day lit hours. It has been shown that the absorption in the *D*-region is inversely proportional to the square of the frequency and thus this region limits the lower range of the useable HF frequencies to rather high values during daytime. Further, hard X-ray (1-10Å) emissions during solar flares can be a significant source of *D*-region ionization in addition to the normal Lyman α (1216Å) radiation and such transient enhancements in electron densities during flares can cause a phenomenon which is popularly known as Short Wave Fadeout (SWF).

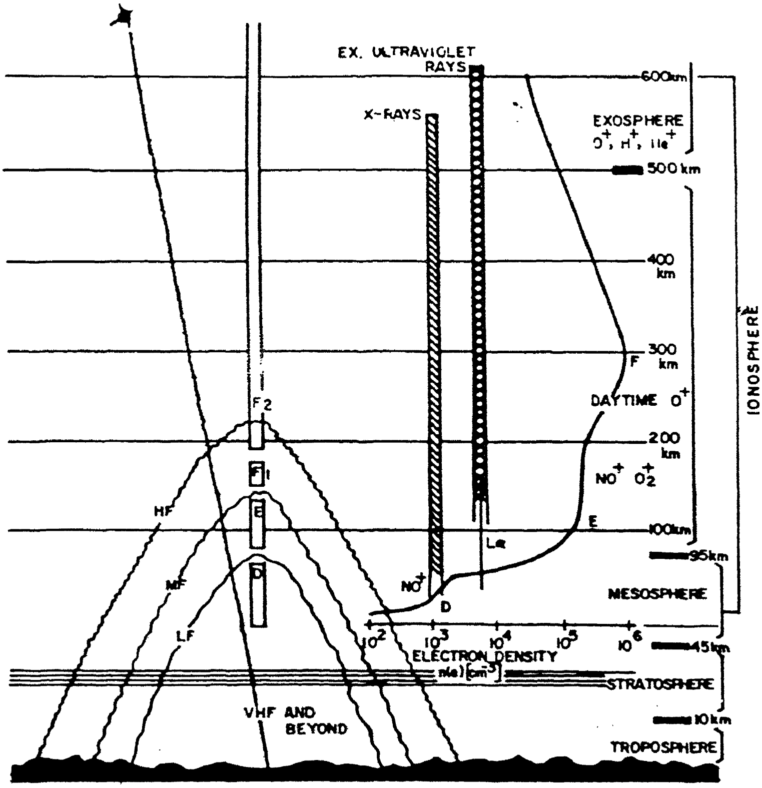


Fig 1 A schematic representation of penetration levels of solar ionizing radiations and daytime ionospheric electron density distribution. Ionospheric Radio Propagation modes in different frequency bands are also depicted.

Dynamical Features of Tropical Ionosphere and their Impact on Radio Systems

The *F*-region of the ionosphere is the most important region for long distance radio communication. In fact, prior to satellite era the only option available for long distance radio communication around the globe was ionosphere-reflected communication in HF bands. It still remains to be so in many of the developing countries like India because of several advantages it has over other modes of communication, the most important being the favourable cost effective factor for HF systems. Another important factor which makes HF communication attractive in tropical countries is the availability of a comparatively wider HF spectrum at these latitudes due to higher electron densities. In addition, the closed field line structure of earth's magnetic field in low latitudes protects these regions from solar particle invasions which can severely disrupt HF communications. This problem is particularly severe at high latitudes where solar energetic particles can find easy access into ionosphere through open field line topology and cause communication disturbances. During certain periods especially in high solar activity years there can be communication failures for days together at high latitudes. Though tropical ionosphere offers several advantages to ionosphere-reflected communication, it also poses some peculiar problems characteristic of these latitudes mainly arising due to the dynamical nature of the ionosphere. The distribution of ionization in *F*-layer is complicated and largely controlled by earth's magnetic field. Several hardly understood processes like electrodynamical, diffusive, chemical, magnetospheric and neutral wind systems come into action and cause a variety of anomalies in electron density distribution in the *F*-layer. In fact, tropical *F*-region is beset with anomalies. The diurnal variation of electron density is not symmetrical with local noon, and there is the well-known Equatorial Ionization Anomaly or Appleton anomaly, the winter anomaly, and the December anomaly etc.

Equatorial Ionization Anomaly

The Equatorial Ionization Anomaly (EIA) is one of the most intensely studied features of tropical and subtropical ionosphere and it refers to the marked departure in the latitudinal distribution of *F*-region ionization from its zenith angle dependence. Instead of an expected ionization maximum around the equator, two distinct crests at about 20 degrees latitude on both sides of magnetic equator manifest during the day[1]. The unique geometry of the geomagnetic field in the sense that it is horizontal over the geomagnetic equator mainly contributes to this anomalous behaviour. The East-West horizontal electric field generated in the E-region over the dip equator due to global scale dynamo action during the day and the earth's horizontal magnetic field which are mutually perpendicular to each other cause upward lifting of the plasma at the equator. The electro-dynamical lifting of plasma over the equator and subsequent plasma diffusion along magnetic field lines to subtropical latitudes known as "Fountain Effect" results in the manifestation of anomalous double hump structure in the *F*-region. The anomaly region is thus marked by steep spatial gradients.

The spatial gradients in the anomaly region can pose a variety of problems to HF radio communications. This can cause asymmetric ray paths, deviation from great-circle propagation, single-hop propagation over long distances without intermediate ground reflection etc.[2,3]. In addition, a peculiar situation of one-way communication can arise because of the large horizontal gradients in this region[4,5]. For example, if the point of reflection of a north-south HF link is poleward of the anomaly peak in the northern hemisphere, the radiowaves incident on the ionosphere for the northern circuit will continuously come across increasing levels of electron density both due to vertical as well as horizontal gradients; while in the north-south path the direction of the horizontal gradient is reversed. Consequently the real Maximum Useable Frequency (MUF) values for the two opposite directions in the same circuit can vary by a large margin (Fig. 2). This situation may result in one-way communication when the operating frequency is close to MUF (predicted ignoring horizontal gradients).

Temporal Gradients in Ionization

The low latitude ionosphere exhibits steep temporal gradients particularly during the dawn and dusk transition periods. Spectacularly large percentage changes in *F*-region critical frequency (f_0F_2) values during sunrise hours have

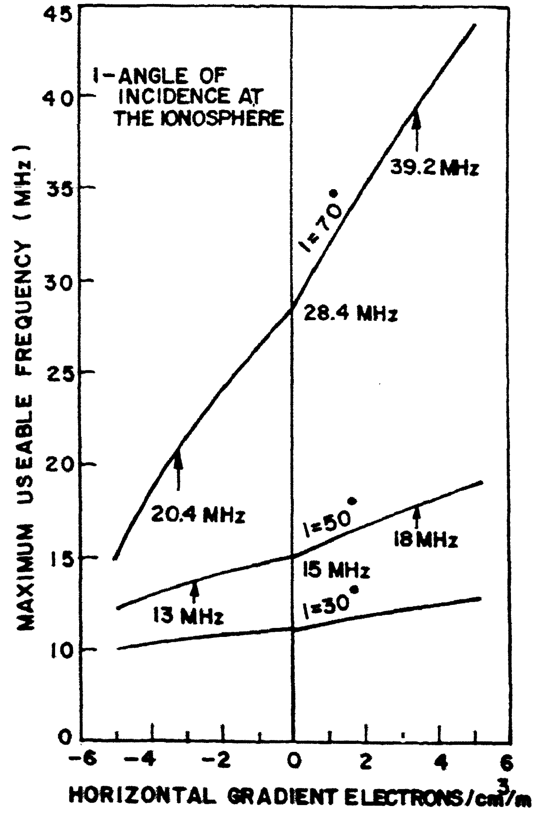


Fig 2 Effects of horizontal electron density gradients on the Maximum Useable Frequencies in HF Communication paths for different angles of incidence (after Lakshmi *et al.*[4]).

been reported for some low latitude stations[4,5]. These steep changes in f_0F_2 values during the sunrise periods will create problems in the case of long distance circuits in the east-west direction involving multi-hop *F*-region propagation because, different *F*-region reflection points fall in the transient location at different periods. Likewise the collapse of the equatorial *F*-region following sunset also can be spectacular, particularly in the Indian longitude sector compared to other sectors[5] (Fig. 3). Such rapid changes can be attributed to the limited capacity of the low latitude magnetic field tubes in holding the plasma. These gradients become less significant with increasing latitude. For high reliability systems it is necessary to have additional frequencies to be used in such transitional periods or even better would be to have adaptive systems discussed in a subsequent section of this paper.

Day-to-day Variability

The day-to-day variability in *F*-region parameters which is apparently unrelated to any specific solar or magnetic event is perhaps the biggest challenge to ionospheric forecasters and this problem is particularly severe in tropical latitudes. There have been several attempts to generate statistical pictures of

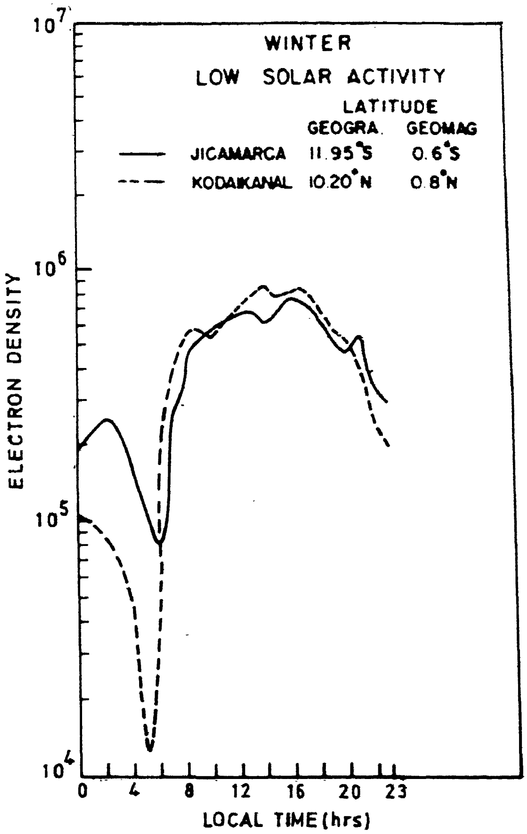


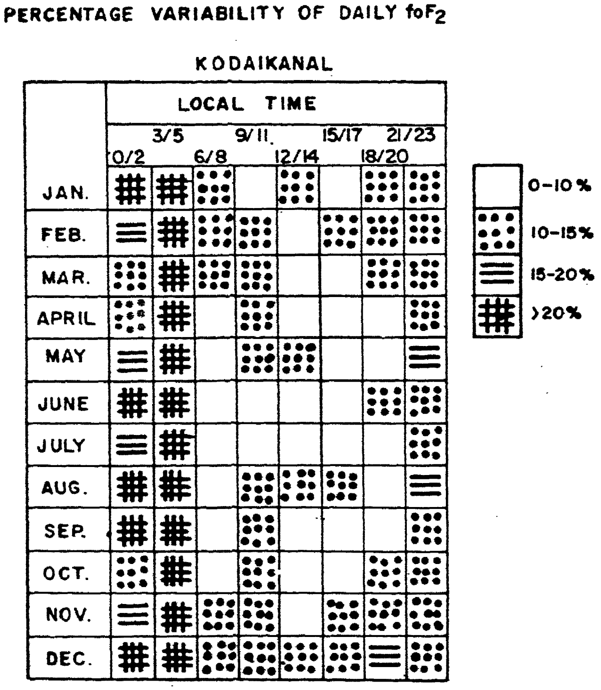
Fig.3 Diurnal plots of *F*-region peak electron densities (Ne) for two equatorial stations Kodaikanal and Jicamarca (after Lakshmi *et al.*[5]).

ionospheric variability using a large volume of data to aid in radio communications[6-8]. The extent of variability is found to be dependent on geographical location, local time, and season (Fig. 4). The day-to-day variability in f_0F_2 values is particularly large during night time in winter and in equinoctial months and in low solar activity years. Extent of variability in f_0F_2 is also found to be large during decaying phase of the solar cycle[7].

A large number of studies conducted using Ionospheric Electron Content measurements (IEC) over India have shown day-to-day variations in IEC values, which are not correlated with solar or magnetic activities[9-12]. Dabas *et al.*[12] reported that the equatorial electrojet has a pronounced influence on IEC at latitudes extending from the equator to as high as 25°N dip latitude and also observed the day-to-day variability to be particularly large around equatorial anomaly peak region (Fig. 5).

It is imperative to have the morphological models of ionospheric day-to-day variability because of its importance in both ionosphere-reflected as well as in trans-ionospheric communications. However, large extents of variability which are often indistinguishable from magnetic storm-time departures in ionospheric parameters at low latitudes render the task of modelling a difficult one.

Large extents of variability necessitates the operating frequencies for HF links to be much lower than the predicted Maximum Useable Frequencies



PERCENTAGE VARIABILITY OF DAILY foF_2

KODAIKANAL

	LOCAL TIME							
	0/2	3/5	6/8	9/11	12/14	15/17	18/20	21/23
JAN.	#	#	⁙		⁙		⁙	⁙
FEB.	≡	#	⁙	⁙		⁙	⁙	⁙
MAR.	⁙	#	⁙	⁙			⁙	⁙
APRIL	⁙	#		⁙				⁙
MAY	≡	#		⁙	⁙			≡
JUNE	#	#					⁙	⁙
JULY	≡	#						⁙
AUG.	#	#		⁙	⁙	⁙		≡
SEP.	#	#		⁙				⁙
OCT.	⁙	#		⁙			⁙	⁙
NOV.	≡	#	⁙	⁙		⁙	⁙	⁙
DEC.	#	#	⁙	⁙	⁙	⁙	≡	⁙

☐	0–10%
⁙	10–15%
≡	15–20%
#	>20%

Fig 4 Ionospheric variability (characterized by coefficient of variation) for Kodaikanal averaged over a solar cycle for different months and local time intervals (after Shastri and Aggarwal[8]).

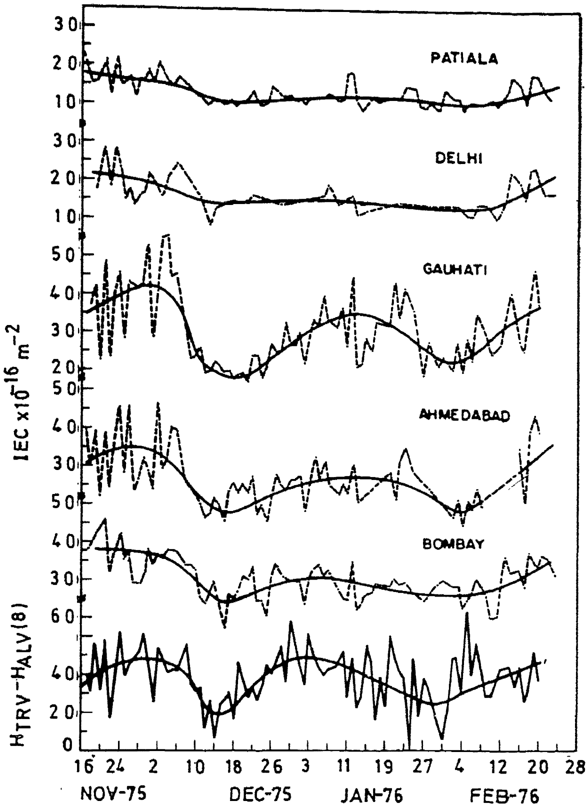


Fig 5 Day-to-day IEC$_{max}$ values from different stations and corresponding electrojet strength during winter 1975-1976 (after Dabas *et al.*[12]).

(MUFs); and at lower frequencies, the power requirements steeply go up resulting in expensive equipment and in avoidable interference problems.

F-Region Irregularities

Ionosphere is not a homogeneous medium and has irregular structures embedded in it. A narrow belt of latitudes around the magnetic equator and another one around auroral latitudes in particular are characterized by the presence of intense ionospheric irregularities. Presence of spread-*F* obvious in ionograms, plume structures in VHF back scatter radar images, scintillations in trans-ionospheric signals can be viewed as signatures of these ionospheric inhomogeneities[13-17]. It is now well established that Rayleigh-Taylor type of plasma instabilities generated in the form of plasma bubbles over magnetic equator initially at the bottom side of the *F*-layer during post-sunset hours develop into spread-*F* and scintillation-producing irregularities in the equatorial and low latitude *F*-region. Under the influence of eastward electric fields the plasma bubbles rise nonlinearly in the *F*-layer as a result of EXB motion producing plasma irregularities along the magnetic field tubes in a wide spectrum of scale sizes on either side of the magnetic equator extending beyond 20° latitude[18-23].

Spread-*F* occurs in the vicinity of magnetic equator usually between 2000-0500hrs and more frequently during high solar activity periods with an equinoctial maxima and is found to be inhibited by magnetic disturbances[24-26]. Some recent studies on ionospheric irregularities using VHF scintillation data have shown that they can be either inhibited or generated during magnetic storms depending on the phase of the storm and their local time of occurrence[27-30]. It is also observed that the very fast fading called the flutter fading in HF radio signals is characteristic of equatorial zone, and occurs especially during post sunset hours when spread-*F* is present. Some salient features of the limitations imposed by these irregularities on satellite-based radio systems are discussed in a subsequent section.

Equatorial Sporadic-E and Radio Communication

Sporadic-*E* (or E_s) is a thin layer with dense patches of ionization around normal *E*-region altitudes. The genesis and properties of sporadic-*E*-layer are markedly different in different zones of the globe. One of the major zones of sporadic-*E* (E_s) ionozation is around geomagnetic equator where the plasma instabilities arising from equatorial electrojet are believed to be responsible for equatorial E_s[31,32]. There have been several excellent studies on equatorial E_s using South American, Nigerian and Indian data and the results are well documented[32-35].

In equatorial latitudes (±6 dip) sporadic-*E*-layer is observed to be present during most of the daytime and electron densities upto a million per cc can occur around local noon hours during equinoctial months. The occurrence probability of such large values is high particularly during solar maximum years (Fig. 6). Several schemes have also been developed for predicting E_s occurrence probabilities in equatorial zone[36,37].

E_s layer is not routinely used for radio wave propagation because of its sporadic nature. However, in certain zones like equatorial zone where E_s exists during most of the day planned communication *via* E_s layer is possible. High levels of sporadic-*E* ionization at *E*-region altitudes yield large operational frequencies (due to obliquity factor) for HF links enabling efficient radio communication over long distances; also the specular nature of E_s reflections contributes to lower transmitter powers. E_s layer can also be detrimental as it might bring in unusually large signal intensities from unwanted transmitters. It can also cause problems to HF circuits by shielding the radio wave from entering the *F*-region. Anomalous long distance TV propagation in VHF bands can also be attributed to presence of E_s[38].

Some Recent Trends in HF System Applications

While the prime user services of HF systems remain to be HF link communications and broadcasting with analog amplitude modulation several new applications have also emerged such as adaptive/digital systems and OTH Radars.

Adaptive HF System

One of the main problems in HF radio communication is the rapid change in the ionospheric characteristics. The operating frequencies need to be

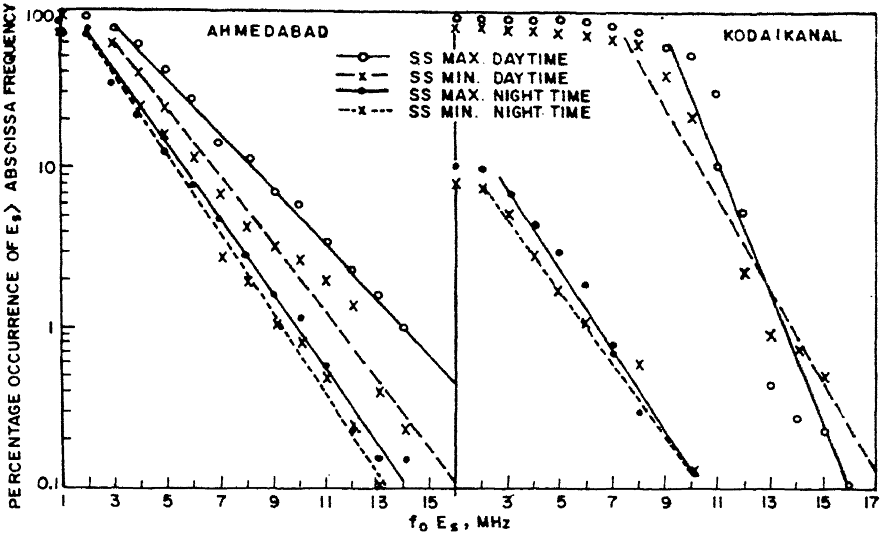


Fig 6 Cumulative distribution of E_s at Ahmedabad and Kodaikanal under various conditions (after Saksena[37]).

changed from time to time to get satisfactory performance. Adaptive systems have the capability for Real-Time-Channel-Evaluation and automatically select the best frequency from the pool of available frequencies. A main pool of frequencies are assigned for a link based on long-term ionospheric predictions and a sub-pool is selected for the day based on short-term forecasts[39]. The main as well as the sub-pools necessarily become large because of the presence of steep ionization gradients in the tropics. One way communication problems that arise because of horizontal gradients in this region and day-to-day variations in ionospheric parameters will only further complicate the situation. High levels of ionospheric absorption in tropical latitudes in turn will increase the transmitter power requirement. The net result would be to exert additional pressure on already crowded HF spectrum in countries like India. The only way to tackle this problem is by generating good prediction models so that proper. optimization of number of frequencies and transmitter power for the link can be made.

Signal fading poses serious problems to data transmission by digital systems. Fading may be caused by several factors such as ionospheric irregularities (spread-*F*), multipath propagation modes including E_s modes, scattered signals from Travelling Ionospheric Disturbances, rotation of axes of polarization, time variations of ionospheric absorption etc.

To minimise the effects due to signal fading one needs to provide proper fade margins while designing the link. This means one must have statistical information on signal variation such as, fade depth, fade rate etc., in addition to morphological aspects of fading occurrence patterns.

OTH Radar

The conventional microwave radars used for ground-based or air-based surveillance have extremely short ranges of the order of a few hundred kilometers and cannot provide adequate advance warning in the contemporary warfare scenario. Several countries in the recent past have successfully commissioned HF Radars on ionosphere – reflected propagation and hence with a range exceeding 10,000 kilometers. These HF radars known as Over-The-Horizon radars have the additional advantage of being less vulnerable compared to satellite-based systems or microwave radars because they can be installed deep within one's own territory. However, the successful operation of these OTH radars is crucially linked to the short-term stability of ionosphere over time periods comparable to the integration times required for unambiguous separation of targets from the overwhelming clutter in the doppler spectrum. *F*-region turbulence and Travelling Ionospheric Disturbances degrade the short-term stability and hence radar performance. Likewise, multipath propagation often due to vertical layer structuring and sporadic *E*-layer cause serious problems. The horizontal ionization gradients produce unpredictable tilts in the propagation path with consequent errors in range and elevation angle and hence in target location[40]. To account for all such variabilities, the short-term prediction of ionospheric parameters assumes special significance so that the integration times can make a compromise between the stability of the ionosphere and the requirement of signal to noise ratio. Adaptive systems with real time ionospheric monitoring should be used for such critical operations.

Ionospheric Models for Radio Communication

Ionospheric Maps

Since early days of ionospheric research, vertical incidence ionograms obtained from ground based ionosonde experiments have been providing information on electron densities but mostly for *E*- and *F*-layers. Information on *D*-region was obtained mainly through rockets and some ground-based techniques such as cross-modulation experiments. With satellites coming into vogue during early 60s high resolution *F*-region information was possible particularly through topside sounder experiments. Several satellite based *in-situ* probes and satellite radio beacon experiments have also added a variety of information on ionospheric parameters. In addition, incoherent scatter radar technique has made possible, the direct measurement of electron density $N(h)$ profiles. These various experimental studies have not only contributed to better understanding of the ionospheric processes but also led to development of reliable ionospheric models. The need for accurate ionospheric models has always existed because of their application potential most importantly in radio communications.

The most comprehensive global model for predicting ionospheric parameters was brought out by CCIR (Report No. 340)[41] using ionosonde data covering a good part of earth's surface. However, large gaps remained in ionosonde coverage in the ocean areas and sparsely populated land areas, and also in the equatorial anomaly region, South Atlantic magnetic anomaly region etc. and numerical mapping techniques were adopted to fill in these gaps[42]. The availability of satellite-based data particularly from topside-sounders has also helped to cover some of these regions. There were several studies to examine the compatibility of ground-based and topside sounder data from Alouette and ISIS satellites[43-45]. Worldwide maps of f_0F_2 have been also produced from Japanese ISS-b satellite data[46-48]. Observations from satellites are particularly valuable for ascertaining sharp, well-defined features in the ionosphere, like those associated with equatorial anomaly region.

Global maps of ionospheric parameters, though are extremely useful in planning global communication, regional maps featuring the local ionospheric characteristics with better resolution have become necessary for improving the prediction accuracy and thereby help in conserving the HF spectrum. The Atlas of Ionospheric Communication parameters over Indian subcontinent brought out by National Physical Laboratory, New Delhi[49] using his resolution data in this region is a significant contribution in this direction. Fig. 7 shows a sample contour map of f_0F_2 (MHz) for the Indian zone for the month of October.

International Reference Ionosphere (IRI-90) models that have come out as a result of sustained efforts from a large number of scientists representing different countries, are the latest and the most comprehensive global models to represent various ionospheric parameters, which include apart from electron density, the electron and ion temperatures and ion composition. These models are expected to serve as standard reference models[50].

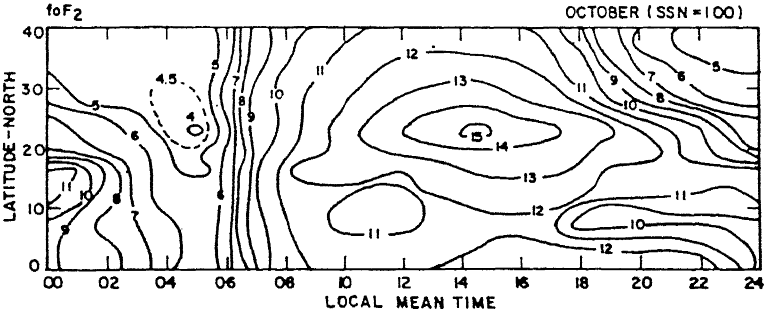


Fig 7 Contour map of f_0F_2 in MHz for Indian zone for October month for sunspot number 100 (after Aggarwal and Shastri[49]).

Ionospheric Indices

The majority of ionospheric prediction models currently in use are based on the excellent long-term correlations that exist between the relevant ionospheric parameters and sunspot numbers[51]. The ideal situation however, would have been to use measured values of solar EUV and X-rays instead of sunspots as an ionospheric index for these are the very radiations which are responsible for producing the ionized layers. Measurement of these radiations is possible only by satellite-borne experiments and some limited data have become available by such experiments only in the recent years. Some preliminary attempts were made using these limited data sets to study solar activity dependent variations in f_0F_2 and IEC values on medium and short-term basis[52-55].

Long series of reliable data on sunspots spanning more than 22 cycles is found to be particularly suitable for long term prediction of ionospheric variations over periods of a few months to several years; however, for predicting short term ionospheric variations especially for periods less than a week, sunspot numbers seem to be of limited use. It was shown that use of EUV flux has certain advantages over sunspots especially during very high sunspot years for predicting ionospheric variation in tropical latitudes[52]. The nonlinear ionospheric variation observed in monthly median f_0F_2 values with sunspot numbers, the so called "Saturation Effect" does not seem to be present with EUV flux (Fig. 8). However, for predicting short-term ionospheric variations, EUV flux does not seem to offer any specific advantage.

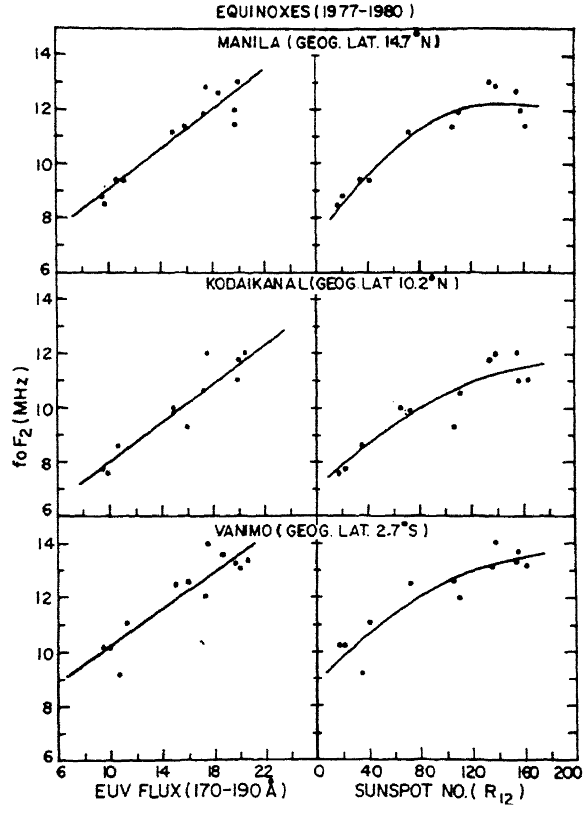


Fig 8 Plots of monthly median noon f_0F_2 vs EUV flux (170-190 A) in arbitrary units and R_{12} for three low latitude stations for equinoxes (after Lakshmi *et al.*[52]).

Disturbed Tropical Ionosphere

The ionosphere in addition to its systematic variations with solar activity, local time, season and latitude also undergoes abrupt and at times random variations lasting for a few minutes to several days. While it is relatively simpler to model the systematic variations, the task of modelling the irregular short-term variations is difficult if not impossible because of their very nature. Short-term prediction models are, however, needed to correct the large errors that are inherent in long term forecasting.

Disturbances in the ionosphere can be usually related to events of solar origin. Geomagnetic storms which occur with a delay of 1 to 3 days consequent to intense clouds of solar energetic particle emissions following solar flares, coronal mass ejections or coronal hole efflux can cause major disturbances in the ionosphere. While geomagnetic storms are delayed by 1 to 3 days after the solar events, there is an additional delay of 10-20 hours in the ionospheric response. The comparatively longer delay times of magnetic disturbances after the solar events enable their prediction and also their ionospheric effects on near-real-time basis. This essentially means development of models (*i*) for predicting a geomagnetic storm based on solar observations and (ii) another for storm-time departures in ionospheric parameters. While the predictability of the storm-inducing solar event itself remains elusive, it is now possible to predict the storms after seeing the solar events through optical, X-ray and radio observations. Daily forecasts on geomagnetic conditions are being issued by the various Regional Warning Centres (RWC) operating under International URSIGRAM and World Day Service (IUWDS) network and these forecasts can be profitably used to predict their likely terrestrial effects including the ionospheric responses.

Ionospheric Storms

Magnetic storm-induced behaviour of ionosphere is highly variable and complex. It is known to be dependent on severity and phase of the storm, occurrence time, latitude, longitude, season etc. There have been several studies on these aspects and the results are well documented[56-59]. The most outstanding features are the dramatic decrease of *F*-region peak densities (N_e) at mid and high latitudes and some modest increase in N_e at low and equatorial latitudes during the main phase of the storm. The *F*-region peak heights are found to increase at all the latitudes.

The observed ionospheric changes arise due to storm-induced changes in dynamics and neutral composition and temperature. Additional current systems set-up during storms can drastically change electro-magnetic drift patterns ($E \times B$) all over the globe. Further, the increased heat input during storms resulting from both precipitating particles and Joule dissipation of ionospheric electric fields at auroral latitudes can bring about changes in neutral temperature and composition. The energy from these high latitude heat sources is transported to low latitudes through gravity waves and or by equatorward winds[60-62]. It is also important to note that no single mechanism can explain all the observed changes, and it is quite likely that several of these mechanisms may simultaneously come into play depending on the latitude and severity of the storm. Molecular-rich equatorward winds can bring about compositional changes (increase in N_2/O) at mid latitudes, thus causing electron density depletion due to increased loss rates, while electrodynamical effects become important for the observed positive responses at equatorial and low latitudes.

MUF Variations during Disturbed Periods

Normally HF circuit planning is based on predicted monthly median values of maximum useable frequencies (MUF) and a knowledge of departures expected in MUFs from their monthly medians during disturbed periods is essential for the operation of any reliable HF link. Exhaustive studies were conducted at National Physical Laboratory, New Delhi using a large volume of data to develop models for prediction of MUF departures in low and equatorial latitudes[63]. These departures in MUFs are dependent on the simultaneous variations that take place in the F-region critical frequencies (f_0F_2) and peak heights (h_mF_2) during disturbed periods. Both these parameters are known to respond positively at these latitudes, however, these positive responses result in compensating effects on MUF. The increase in f_0F_2 will increase the MUF values while the increase in h_mF_2 will tend to decrease the values for the circuit. It should be mentioned here that the changes in h_mF_2 become important for circuits greater than 1000km and for shorter circuits the variations in MUF are essentially decided by variation in f_0F_2. In the mid and high latitudes the combination of negative responses in f_0F_2 and positive h_mF_2 responses would result in large negative deviations in MUFs leading to serious disruptions in HF networks. In comparison the storm-time variations in MUFs at low latitudes are modest and normally of positive kind. This is one of the reasons that makes HF communications very attractive at low latitudes.

Disturbed-time models developed at NPL, New Delhi indicate that the ionospheric responses to magnetic disturbances in Indian Zone vary widely starting from positive responses in 8-20°N geographic latitude belt to predominantly negative response around 30°N (Figs. 9 and 10) interspersed with a very complex situation of both positive and negative responses at latitudes around the peak of the equatorial anomaly. On some occasions specially during se-

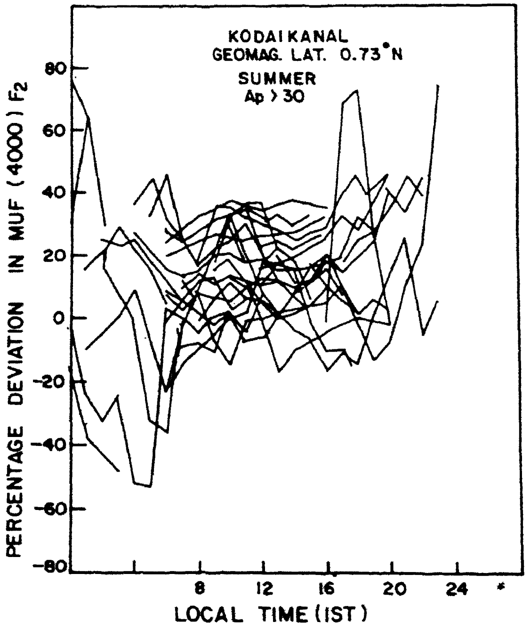


Fig 9 Percentage deviations in MUF (4000)F_2 from monthly median values for several disturbed days during summer for Kodaikanal (after Lakshmi *et al.*[63]).

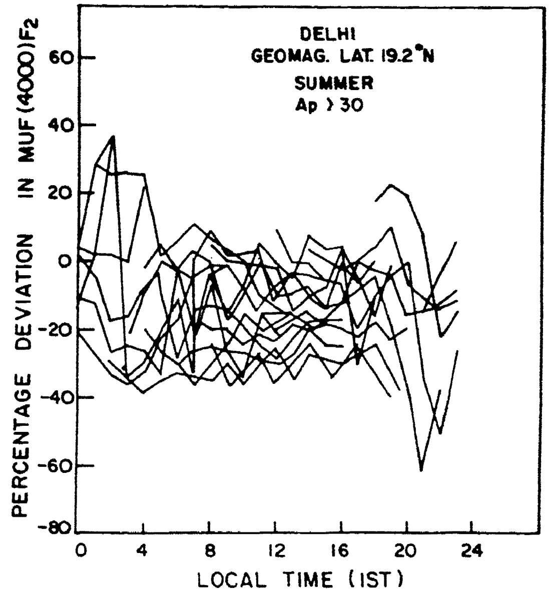


Fig 10 Percentage deviations in MUF $(4000)F_2$ from monthly median values for several disturbed days during summer for Delhi.

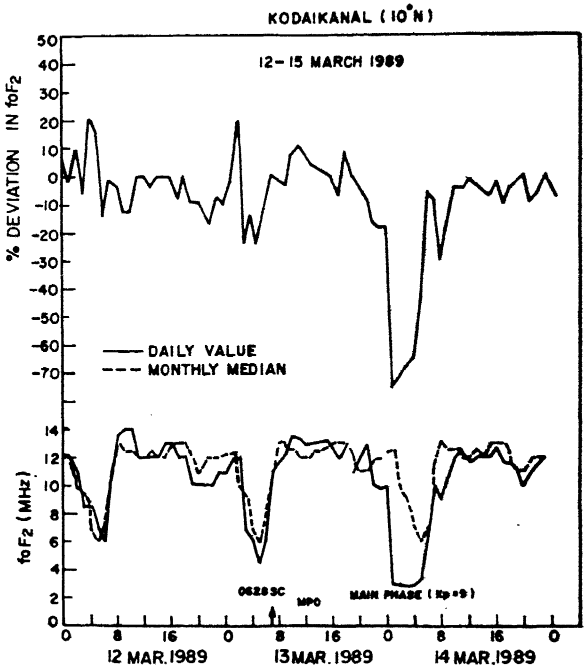


Fig 11 Shows daily values of f_0F_2 at Kodaikanal during the great storm of 13 March 1989 and their percentage deviations from the monthly medians. Spectacularly large negative deviations during the main phase are to be noted.

vere storms, ionospheric responses during night at equatorial latitudes were observed to be negative and significantly large[64,65] and it is these situations which pose serious problems to radio communications. Fig. 11 shows the ionospheric response for a very severe storm that occurred in March 1989 during which the Kp touched the maximum possible plunging the f_0F_2 to as low as 80% below the median value. It may be noted that such large negative responses are exceptional at low latitudes and are restricted to severest of the storms. Exclusive models are needed for predicting the ionospheric departures at equatorial and low latitudes during such severe magnetic storms.

Trans-Ionospheric Propagation Effects

The trans-ionospheric propagation effects were first noticed several decades ago as they imposed serious limitations at certain times on Radio Astronomy studies in the VHF. Subsequently, modelling, quantification and compensation for these effects became critical with the evolution of a variety of satellite-based radio systems. While a detailed treatment is beyond the scope of this article, salient features of the limitations imposed by the *F*-region irregularities as well as the total electron content in the ionosphere are discussed briefly below:

Ionospheric Scintillations

The ionosphere, *F*-region in particular, is imbedded with electron density irregularities of all scale sizes; and these irregularities are constantly in motion. When radio waves pass through this train of moving irregularities, the radio waves are subjected to fluctuations in amplitude, phase, polarisation and angle of arrival and produce the so called scintillations at the receiver. Until the sixties, most of these studies were limited to amplitude scintillations in the VHF band. By extrapolating the frequency law observed in the VHF band, it was quietly assumed that scintillations in the GHz band would be negligible and would not affect satellite systems in the microwave region. However, it was realised in early seventies that very strong scintillations[66] can occur even at GHz frequencies, especially in the tropical belt. Frequency dependence studies[67] subsequently have confirmed that the dependence becomes weak during strong scintillations and the frequency dependence practically disappears when scintillations saturate with Rayleigh type of fading[68].

There are two major interests in the study of scintillations. One is to prepare reliable models based on the extensive studies of scintillation morphology to plan operation of space-based radio systems for communications, control, command and intelligence and several other emerging areas of applications upto several GHz. The other interest of fundamental nature is to use scintillation experiments as diagnostic tools or tracers to understand the physics of the ionosphere and the dynamics of the irregularity structures. One such recent experiment in India[69] has yielded scientifically rewarding information on equatorial plasma bubble dynamics. The so-called bubbles are known to be generated through Rayleigh-Taylor instabilities following the post-sunset sudden collapse of the *E* and the lower *F*-region. The bubble grows and non-linearly driven up by polarization EXB forces and simultaneously drifts westward relative to the ambient ionosphere.

The plasma bubble rise velocities were estimated from VHF radar observations[14], from satellite observations[70], from spaced ionosonde measurements[71] and in the Indian zone from multistation scintillation measurements[72]. In view of the unique information derived from the recent campaign[69] conducted in India, a brief description is in order. The 4GHz telemetry signals from INSAT-1B (74°E) and INSAT-1C (94°E) were recorded simultaneously at Sikanderabad (24.48°N and 77.1°E, Dip 42°) near Delhi. The subionospheric points for the two geostationary satellites are separated by 121 km in the east-west direction. Fig. 12 shows a sample record depicting the characteristics and time delay in the onset of scintillations between the two satellite signals. The east-west plasma bubble velocities derived from such records on a number of occasions vary from 28 to 200m/sec with a median of 75m/sec. Similar experiment was conducted at a low latitude station, Chenglepet (12.07°N and 79.92°E, Dip 10.5°) using the same two satellites. The mean eastward velocities derived from Sikanderabad and Chenglepet are shown in Fig. 13. Also shown are the velocities derived from Jicamarca radar and from a spaced VHF polarimeter over Brazil. A number of such planned campaigns are necessary to understand the dynamics of ionospheric plasma which is driven by the neutral wind and electrodynamical forces in the presence of restrictions imposed by the terrestrial magnetic field.

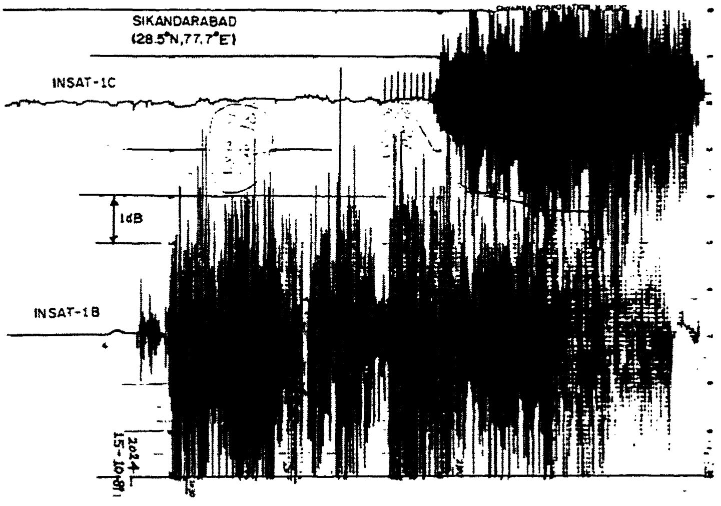


Fig 12 A sample record showing the time delay in the onset of scintillations on INSAT-1B and INSAT-1C telemetry signals at Sikanderabad (after Dabas *et al.*[69]).

The applications' interest of course is overwhelming as is obvious from the poor video received from satellites often even at the low rate of information currently employed. The information pulse length will decrease (and signal bandwidth will increase) as the rate of information transmission increases. Consequently, the multipath spread which is small and relatively insignificant for slow speed systems or analog systems will cause problems for the high speed modulation rates planned in future satellite systems. Good morphologi-

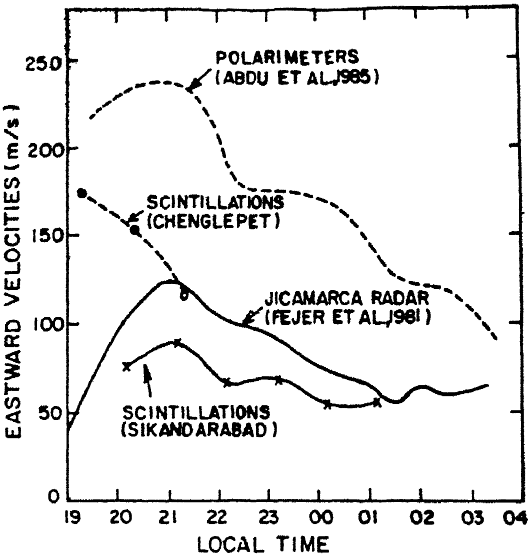


Fig 13 Mean eastward velocities determined over Sikanderabad (crosses) and Chenglepet (open circles) as a function of local time. Jicamarca radar measurements (solid line) and spaced VHF Polarimeter measurements over Brazil (dashed line) are also shown (after Dabas *et al.*[69]).

cal and dynamical models of scientillations occurrence are necessary to restrict the transmission rates to levels with tolerable bit error rates.

Effects due to the Total Electron Content of the Ionosphere

The radio wave, as it traverses through the ionosphere, is subjected to group delay, RF carrier phase advance, Doppler shift of the RF carrier, Faraday rotation of the plane of polarisation and distortion of transmitted pulses. Perhaps the most problematic of all is the time delay which is proportional to the total electron content and inversely proportional to the square of the operating frequency. Even at a frequency of 1GHz, the time delay can result in a range error exceeding 40 metres in satellite based ranging systems. While it is possible to employ a two-frequency time delay correlation, in several systems it is not possible to have such a large band width available in the navigation or ranging systems. This makes it essential to model the electron content and use it for correcting the time delay error. With present models, a 50% correction is possible; better models can extend this correction to at least 75 per cent.

Conclusion

The ionosphere located in our near geo-environment offers us a very inexpensive mode of communicating in the HF spectrum especially in the tropics. However, the tropical ionosphere is also a debilitating factor with intense irregularities which imposes serious limitations on high reliability satellite-based radio systems. With the sophistication in terminal equipment and associated computers, the uncertainities in the propagating media set the limits for reliability in contemporary radio systems. Thus, characterising the media including microstructure and transient response to solar and magnetic events is essential to meet our expanding communication requirements. Experiments should also be properly devised to understand the plasma physics and electrodynamics of this region so that accurate prediction models could be developed to optimally utilise this unique natural resource.

References

1 R G Rastogi *J geophys Res* **64**(1959) 727

2 J O Oyinloye *Handbook on Radio Propagation for Tropical and Subtropical Countries* (Eds A P Mitra, B M Reddy, S M Radicella, J O Oyinloye and S Feng) URSI Committee on Developing countries UNESCO subvention 1984-1985 (1987)27-58

3 *CCIR Recommendations and Reports of CCIR* International Telecommunication Union Geneva **Rep 250-6** (1986)

4 D R Lakshmi, S Aggarwal, P K Pasricha and B M Reddy *Indian J Radio Space Phys* **8**(1979) 101

5 D R Lakshmi, S Aggarwal and B M Reddy Low latitude aeronomical processes (Advances in Spaces Exploration) *In: COSPAR Sym S* **8** [Ed A P Mitra] Pergamon Press, Oxford (1980) 141-144

6 S Aggarwal, D R Lakshmi and B M Reddy *Solar Terrestrial Predictions* (*Proc*) (Ed R F Donnelly) US Deptt of Commerce **1**(1979) 134-139

7 S Aggarwal *Indian J Radio Space Phys* **14**(1985) 73

8 S Shastri and S Aggarwal *Mem G S I* No 26 Geological Society of India Bangalore (1992) 483

9 T R Tyagi and Y V Somayajulu *Radio Sci* **1**(1966) 1125

10 Malkiat Singh and H S Gurm *Indian J Radio Space Phys* **8**(1979) 306

11 P V S Rama Rao, D Nru and M Srirama Rao *Proc COSPAR/URSI Symp* [Ed A W Wernik] Polish Scientific Publishers Warsaw (1981)175-187

12 R S Dabas, P K Bhuyan, T R Tyagi, R K Bhardwaj and J B Lal *Radio Sci* **19**(1984) 749

13 D T Farley, B B Balsley, R F Woodman and J P Mcclure *J geophys Res* **75**(1970) 7199

14 R F Woodman and C Lahoz *J geophys Res* **81**(1976) 5447

15 R G Rastogi and R F Woodman *J atmos terr Phys* **40**(1978) 485

16 S Basu, Su Basu, J Aarons, J P Mcclure and M D Cousins *J geophys Res* **83**(1978) 4219

17 S Basu, J P Mcclure, Su Basu, W B Hanson and J Aarons *J geophys Res* **85**(1980) 5119

18 R T Tsunoda *J atmos terr Phys* **42**(1980) 743

19 R T Tsunoda *J geophys Res* **86**(1981) 139

20 B G Fejer and M C Kelley *Rev Geophys Space Phys* **18**(1980) 401

21 L S Ossakaw *J atmos terr Phys* **43**(1981) 437

22 Y V Somayajulu, S C Garg, R S Dabas, L Singh, T R Tyagi, B Lokanadham, S Ramakrishna and G Navneeth *Radio Sci* **19**(1984) 707

23 R S Dabas and B M Reddy *Radio Sci* **21**(1986) 453

24 H Chandra and R G Rastogi *Ann Geophys* **28**(1972) 37

25 R G Rastogi *J atmos terr Phys* **42**(1980) 593

26 M A Abdu, I S Batista and J A Bittencourt *J geophys Res* **86**(1981) 6836

27 A Dasgupta, A Maitra and S Basu *Radio Sci* **16**(1981) 1455

28 R S Dabas, D R Lakshmi, B M Reddy *Radio Sci* **24**(1989) 563

29 J Aarons *Radio Sci* **26**(1991) 1131

30 D R Lakshmi, R S Dabas, B C N Rao and B M Reddy *Radio Sci* **28**(1993) 389

31 D T Farley *J geophys Res* **63**(1963) 6083

32 D T Farley *J atmos terr Phys* **47**(1985) 729

33 P Bandyopadhay and H Montes *J geophys Res* **68**(1963) 2453

34 J O Oyinloye *Radio Sci* **4**(1969) 765

35 A E Giraldez, L Puig and T Lama *Geoacta* **XI**(1981)

36 C E Agu and C A Onwumechili *J atmos terr Phys* **43**(1981) 809

37 R C Saksena *Indian J Radio Space Phys* **4**(1975) 318

38 R C Saksena *Indian J Radio Space Phys* **8**(1979) 351

39 D Borgmann *Solar-Terrestrial Predictions Proc Workshop Leura Australia* Oct. 16-20, 1989 (Eds R J Thompson *et al.*) National Oceanic and Atmospheric Administration Environmental Research Laboratory Boulder USA **1**(1990) 206-214

40 B M Reddy Symp Dig *Skywave Over-the-Horizon Backscatter Radar Concepts and Techniques* 11-12 Oct. 1988 Electronics & Radar Development Establishment Bangalore (1988) 13-14

41 *CCIR Atlas of Ionospheric Characteristics* International Telecommunication Union Geneva Rep. **340**(1970) Revised (1983)

42 W B Jones and R M Gallet *Telecommun J* **29** (1965) 18

43 R Neshizaki and M Nagayama *J Radio Res Lab (Japan)* **16** (1969) 227

44 J D Burge, J W King and A J Slater *Telecommun J* **40**(1973) 356

45 W B Jones and P G Stewart *Radio Sci* **5**(1970) 773

46 N Matuura *Atlas of Ionospheric Critical Frequency* (f_0F_2) *obtained from Ionosphere Sounding Satellite-b Observations* Part 1 *Radio Research Labs* Ministry of Posts and Telecommunications, Japan (1979)

47 N Matuura *Atlas of Ionospheric Critical Frequency* (f_0F_2) *Obtained from Ionosphere Sounding Satellite-b Observations* Part 2 *Radio Research Labs* Ministry of Posts and Telecommunications Japan (1980)

48 N Matuura *Atlas of Ionospheric Critical Frequency* (f_0F_2) *obtained from Ionosphere Sounding Satellite-b Observations* Part-3 *Radio Research Labs* Ministry of Posts and Telecommunications Japan (1981)

49 S Aggarwal and S Shastri *An Atlas of Ionospheric Communication Parameters* National Physical Laboratory, New Delhi (1991)

50 D Bilitza *International Reference Ionosphere* 1990 *URSI/COSPAR National Space Science Data Center/World Data Center–A Rockets and Satellites* 90-22 NASA Maryland USA (1990)

51 B M Reddy, S Aggarwal, D R Lakshmi, S Shastri and A P Mitra *Solar Terrestrial Proceedings* (Eds R F Donnelly) U S Dept of Commerce USA **1**(1979) 118-133

52 D R Lakshmi, R S Dabas and B M Reddy *J atmos terr Phys* **50**(1988) 207

53 R P Kane *J atmos terr Phys* **54**(1992) 463

54 N Balam, C J Bailey and Jayachandran *Planet Space Sci* **41**(1993) 141

55 R S Dabas, D R Lakshmi and B M Reddy *Pure appl Geophys* (1993) (*in Press*)

56 S Matsushita *J Geophys Res* **64**(1959) 305

57 K I Maeda and T Sato *Proc Inst Radio Engrs Aust* **47**(1959) 232

58 Y V Somayajulu *J geophys Res* **68** (1963) 1899

59 G Rajaram and R G Rastogi *Ann Geophys* **25**(1969) 795

60 H Rishbeth *J atmos terr Phys* **37**(1975) 1055

61 G W Prolss *J geophys Res* **86**(1981) 2350

62 M Blanc and A D Richmond *J geophys Res* **85**(1980) 1669

63 D R Lakshmi, B M Reddy and S Shastri *Indian J Radio Space Phys* **12**(1983) 118

64 K M Kotadia *J atmos terr Phys* **24**(1962) 975

65 D R Lakshmi, B C N Rao, A R Jain, M K Goel and B M Reddy *Ann Geophys* **9** (1991) 286.

66 R R Taur *Ionospheric Scintillation at* 4 *and* 6 *GHz COMSAT Tech Rev* (1973) 1

67 E J Freemouw and C L Rino *Modelling of Transionospheric Radio Propagation RADC-TR-*76-35 Rome Air Development Center Griffiss AFB, NY(1976)

68 J Aarons and S Basu *Handbk of Geophysics and the Space Environment Section* 10.7 (Ed A J Jursa) Air Force Geophysical Lab USA (1985)

69 R S Dabas, P K Banerjee, S Bhattacharya, B M Reddy and S Singh *J atmos terr Phys* **54**(1992) 893

70 J P McClure, W B Hanson and R A Hoffman *J geophys Res* **82** (1977) 2650

71 M A Abdu, R T De Medeires, J H A Sobral and J A Bittencourt *J geophys Res* **88**(1983) 9197

72 R S Dabas and B M Reddy *Radio Sci* **25**(1990) 125

THE AURORAL IONOSPHERE—A TELEVISION SCREEN FOR PROCESSES IN THE DISTANT MAGNETOSPHERE

GIRIJA RAJARAM

Indian Institute of Geomagnetism, Colaba, Bombay 400 005, *India*

(*Received* 16 *August* 1993; *Accepted* 19 *August* 1993)

Introduction

The Earth's high latitude regions form a locale where solar-terrestrial energy exchange processes operate with full vigour, and leave visible signatures in the near-geospace environment. The optical manifestation of the energy input is the *aurora borealis* in the northern hemisphere, and the *aurora australis* in the southern hemisphere. *Aurora* is the name given to the shimmering, iridescent, hues of pale violet, green, or pink, which are seen in the night sky, generally at latitudes between 65° and 75° on Earth. The luminescent structures can take on various shapes, such as bands, rays, arcs or even parallel curtains in the sky[1] and are understood to occur at altitudes 100-1000 km in the region known as the bottomside and topside *ionosphere.* Shown in Fig. 1 are some examples of the various dramatic forms which the aurora can assume[2].

The aurora manifests itself as optical phenomena, and is really a pointer to marked electromagnetic changes taking place in the Earth's space environment. These changes originate in deep geospace at distances of some hundreds of Earth radii (1 Earth radius = 1 *Re* = 6370 km) away. In other words, the gaseous upper atmosphere of Earth at 100-1000 km altitude, serves as a screen on which are projected the energetic plasma processes at work in the distant magnetosphere.

Characteristics of the Aurora

When viewed globally from satellites, it becomes clear that the aurora generally occurs over oval-shaped regions, around the magnetic dip poles of Earth in the north and the south; the ovals are pulled away from the dip poles on the nightside. The diameter of the oval is about 3000 km (14° lat on either side of the pole) during magnetically Quiet times, but grows to about 6000 km (28° lat on either side of the pole), during Disturbed times—this points to an equatorward movement of aurora during disturbances in the Earth's electromagnetic environment. While the lower edge of an auroral curtain is invariably at about 100 km, the upper edge can be anywhere between 400-1000 km altitude. An auroral curtain can spread over a few thousand kilometers in the east-west direction, but the thickness of the curtain seldom exceeds 1 km. Aurora basically are of two types—the "diffuse" type and the "discrete" type. The former refers to faint glows in the auroral skies, while the latter manifests as brighter curtain-like, structured displays of light. "Discrete" aurorae are again sub-divided into "faint"

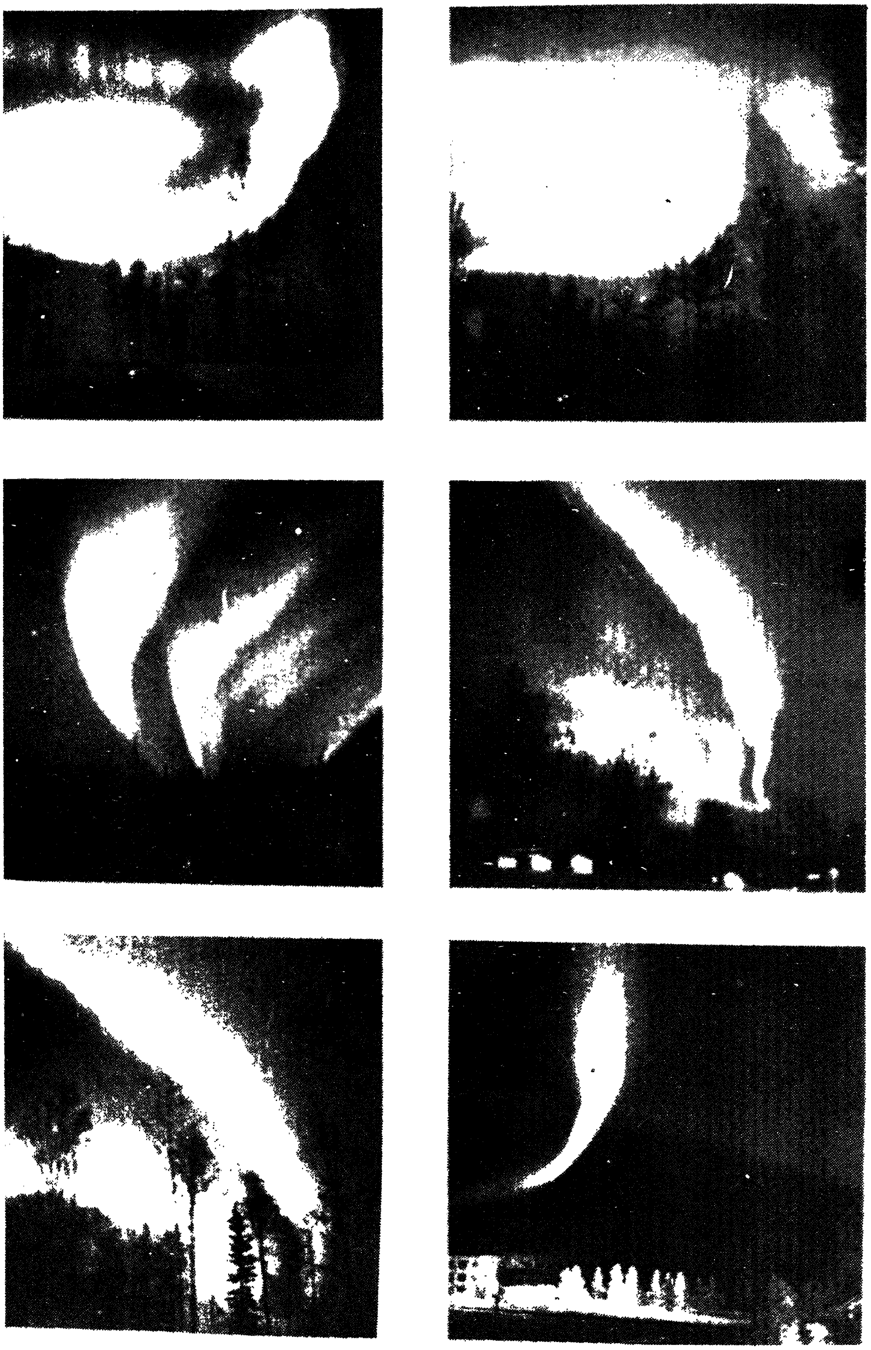


Fig 1 The aurora can take on various shapes and structures ranging from parallel curtains in the sky to deep folds in a single curtain. This figure shows some of the forms which it can attain[2]. Its luminescence in the dark, winter, northern high-latitude skies, earned the phenomenon the name "aurora borealis" i.e. "northern dawn"

curtains and "bright" curtains. A composite of the "discrete" and "diffuse" auroral forms is shown in Fig. 2 after Akasofu and Kamide[3].

Each of the features of the aurora discussed above has a story to tell about physical processes at work in different regions of the distant magnetosphere. Deciphering this information from auroral electromagnetic signatures, is currently an area of active research interest in the space sciences[4].

The phenomenon of aurora basically occurs because 1) the Earth has a gaseous atmosphere, 2) the Earth has a magnetic field, and 3) energetic charged particles from the Sun enter the Earth's environment. In order to understand the auroral phenomenon, it is necessary to know something about the different regions of geospace.

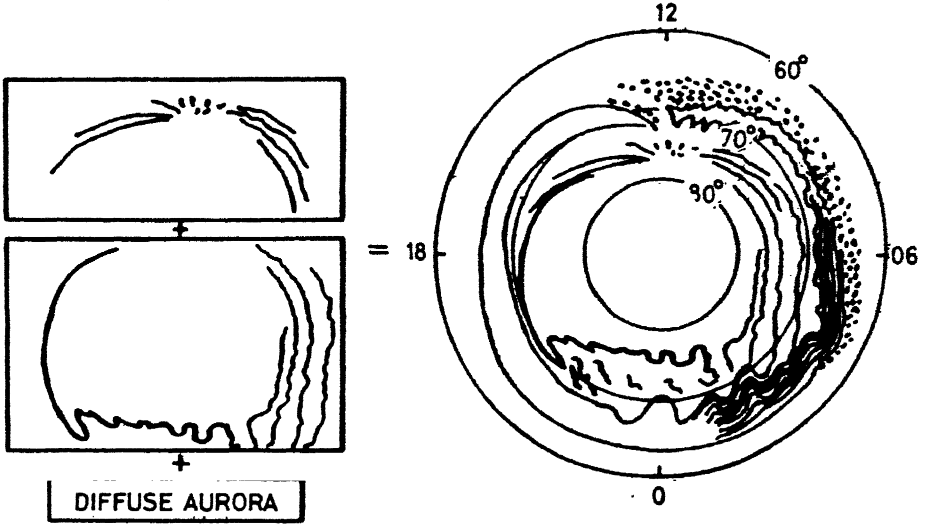


Fig 2 A composite global picture of "discrete aurora" (faint curtains in the dayside, bright curtains in the nightside), and "diffuse aurora" (a faint glow at lower latitudes of the auroral zone at all hours)[3]

The Geomagnetosphere

The Earth like the Sun and other planets of the solar system (indeed for that matter many celestial bodies), has a magnetic field. This magnetic field (geomagnetic field) is essentially dipolar in nature (higher order components are much weaker especially at distances of a few Earth radii), and it extends out into geospace in a 3-dimensional form. The Sun emits solar wind, which consists of electrons and protons with energy of a few electron Volts, together with a few nuclei of helium and calcium. If there was none of this low-energy charged particle emission from the Sun the geomagnetic field would be almost symmetrical in space, with a closed field line configuration all around as shown in Fig. 3(*a*). With the presence of solar wind, however, the solar particles exert a dynamic pressure when they encounter the Earth's magnetic force field. It is interesting to note that during Quiet solar conditions this solar wind velocity of 400 km/sec is about 15 times the sound velocity, and about 6 times the hydromag-

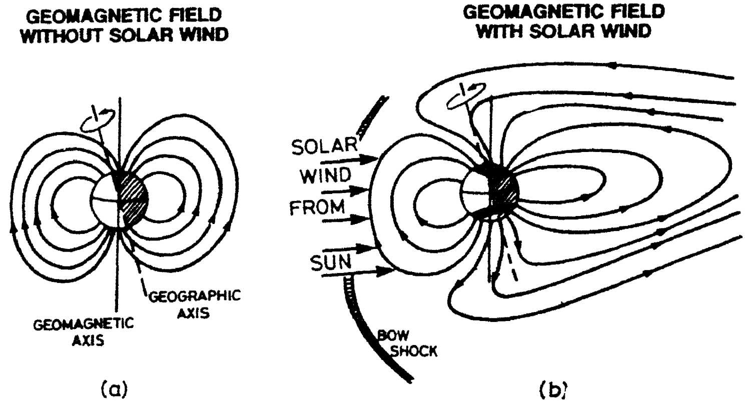

GEOMAGNETIC FIELD WITHOUT SOLAR WIND
GEOMAGNETIC FIELD WITH SOLAR WIND
SOLAR WIND FROM SUN
GEOGRAPHIC AXIS
GEOMAGNETIC AXIS
BOW SHOCK
(a)
(b)

Fig 3 A dipole magnetic field would extend symmetrically into space in all directions as in Fig. 3(*a*). Solar wind compresses it on the dayside, and draws it out into a long tail on the nightside, sometimes as far as 1000 Earth radii. A Bow shock is formed upstream on the dayside because the solar wind senses the geomagnetic field well before it reaches the latter.

netic wave velocity, all in the interplanetary medium[5]. The geomagnetic field resists the dynamic pressure of solar wind with a magnetic pressure, and a state of equilibrium occurs when the dynamic and magnetic pressures balance. This equilibrium condition is expressed by

$$\frac{1}{2}\rho V^2 = \frac{B^2}{8\pi},$$

where ρ—density of solar wind in Number cm^3

V—velocity of solar wind in km/sec

and B—intensity of Earth's magnetic field in Gauss at the boundary where equilibrium is attained.

Strictly speaking this relationship applies only to two-dimensional magnetic fields[6] but it is often used in astrophysics for steady state three-dimensional magnetic fields.

This equilibrium between dynamic and magnetic forces results in the Earth's magnetic field being compressed on the dayside, and being drawn out into a long tail on the nightside. The net result is a teardrop-shaped structure, known as the *Geomagnetosphere*; it is shown in Fig. 3*b* and its boundary is called the magnetopause. On the sunward side, the magnetopause seems to form at distances of 10-14 Re [Coleman *et al.*[7] from the PIONEER 5 measurements], while on the Dawn-Dusk sides, it forms at 30-40 Re[8,9] [from the EXPLORER 10 observations].

The two arcs shown in Fig. 3(*b*) in the northern and southern hemispheres of Earth are the "*auroral regions*" and when mapped outwards along geomagnetic field lines, they project on to the last set of closed magnetic field lines both on the dayside and on the nightside. This set of field lines as we shall see later, plays a vital role in accumulating and accelerating charged particles of solar wind origin into the auroral upper atmosphere known as the "auroral ionosphere', and hence causing the phenomenon of aurora.

The magnetosphere has a rather complex, layered structure, in which different regions are characterised by plasma populations with different energies, and physical processes operating at different energy levels[10]. In Fig. 4 is shown a polar cross sectional cut of a 3-dimensional magnetosphere with an equatorial girdle called the *Low Latitude Boundary Layer.* Regions which are relevant to the formation of aurora are labelled in this figure, and will be referred to in the sections that follow.

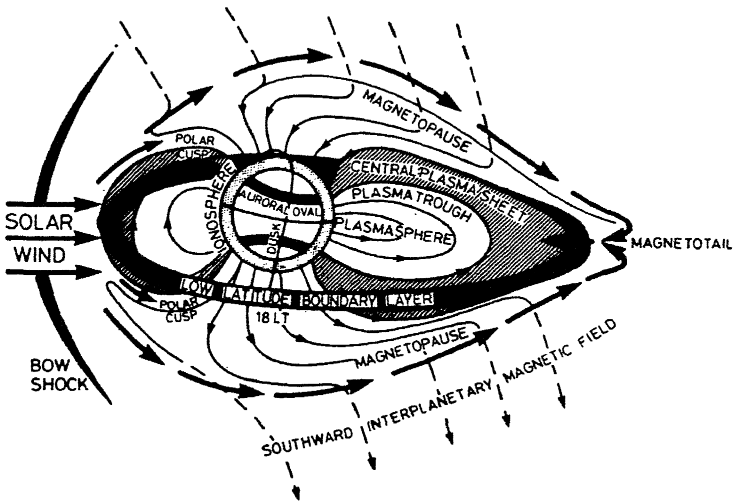


Fig 4 A cross-sectional view of the magnetosphere with its many-layered structure. The solar wind flows across the IMF-geomagnetic field linkages, and acts as a huge power generator. The Low Latitude Boundary Layer lies just within the Magnetopause, and sits over the equatorial magnetosphere like a girdle.

Solar Emissions

The Sun, our nearest star, is understood to be a massive sphere of hydrogen gas, in which protons (hydrogen ions) are continuously being converted into helium nuclei, through the process of nuclear fusion. The temperature of its visible surface (the photosphere) is about 6000°K, but above this it rapidly increases to more than a million degrees Kelvin in the *corona,* which is the far-reaching tenuous atmosphere of the Sun. These high temperature induce a continuous "boiling away" of solar matter into space all around the Sun, in the form of low and medium energy (eV and keV) electrons and protons. This continuous emission was referred to as the "solar wind" in the second section.

The Sun also has a dipolar magnetic field which because of the star's rotation, is stretched out in the solar equatorial plane to form spiral lines of magnetic force (cf. Merrill and McElhinny[11]). This field is known as the Interplanetary Magnetic Field (IMF) because it pervades the entire space between the planets of the solar system. The IMF is also known to have a "sector" structure, with the magnetic field lines alternately pointing away from and towards the Sun[12]. In the case of Earth, this direction of the IMF crucially decides the transfer of energy from the solar wind to the geomagnetosphere[13]. Thus the field and particle emissions of the Sun—the IMF and the solar wind respectively—stretch out deep into interplanetary space, and any planet which encounters these, finds its space-environment altered. Fig. 3(*b*) is a typical example of how these emissions affect planets with a magnetic field. The Sun also emits in other regions of the entire electromagnetic spectrum, namely the X-ray, gamma ray, optical microwave and radio wave etc., but we will not elaborate on these here. We shall instead concentrate on the effects caused by the solar wind and IMF on the Earth's environment.

The Solar-Terrestrial Field and Particle Interaction

The Dayside Aurora

The first obstacle which the solar wind encounters as it approaches Earth is the *Bow Shock*. The Bow Shock is formed slightly upstream of the last closed field line of Earth[14,15], much as a sonic shock is formed by a spherical obstacle in a wind-tunnel. The shock serves to somewhat check the speed of the solar wind, and to introduce a measure of disorder in its progress. The next barrier is the *Magnetopause* constituted by the boundary field lines on the dayside, and the solar wind is constrained to flow over and around it. On the dayside however, there is a Cusp-shaped arc in both hemispheres (labelled as Polar Cusp in Fig. 4), which represents the transition from the closed field line to the open magnetic field line structure. At these regions there is magnetic neutrality, and the solar wind is able to penetrate freely till it meets the Earth's ionosphere. Here the solar wind particles give up their energy to the atoms and molecules of O_2, N_2 and NO which constitute the Earth's upper atmosphere, through reactions which will be described later. This causes the atmospheric gases to get ionised and to change their energy levels; they emit optical light when returning to lower energy levels. This constitutes the *dayside aurora*, and the auroral emissions of the 'discrete' type here are relatively weaker than those on the nightside. This is because the solar wind particles enter the magnetically neutral regions of the polar cusps without undergoing acceleration unlike in the magnetotail where electromagnetic acceleration is at work. Shown in Fig. 5 is a satellite photograph of "discrete" aurora[16]. Very noticeable in this picture is the large gap in the noonside, created by the absence of energetic processes which give rise to the "discrete" aurora.

The Nightside Aurora

We now follow in Fig. 4, the solar wind particles which have not entered the polar cusps, but have continued their flow along the outer surface of the magnetopause towards the Earth's nightside. These particles turn around at the end of the magnetotail because here they encounter another region of magnetic neutrality known as the *Neutral Sheet*. The neutral sheet essentially separates oppositely-directed lines of magnetic field which are drawn far out into space on the nightside, as far as a few hundreds of Re, by the flow of solar wind. Once the

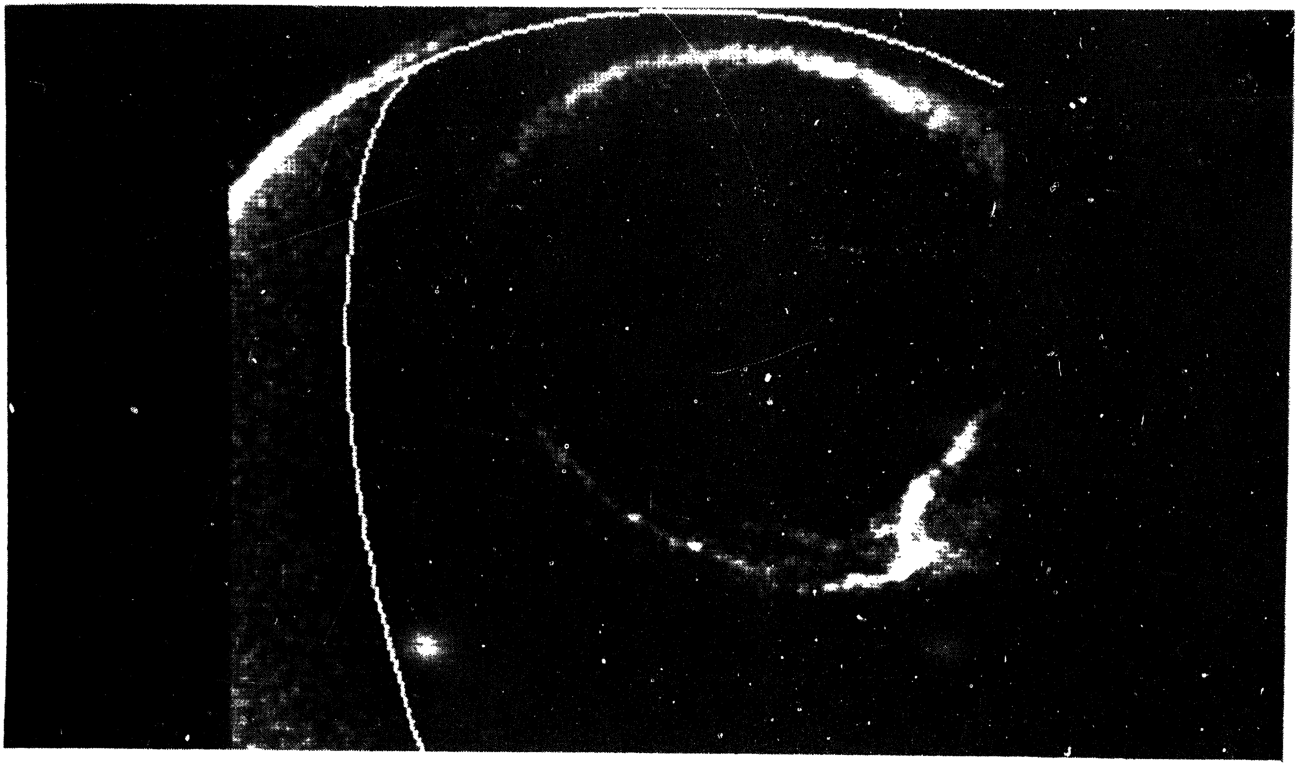


Fig 5 The local noon sector generally experiences much less of the bright curtains which mark "discrete" aurora, and which are present in other local time sectors, most so in the night sector. This points to lack of acceleration processes for electrons in the dayside. The gap in "discrete" aurora at noontime is clearly seen in this satellite image of the auroral oval taken on 8 Dec 1981 at 0154 UT[16]. The Earth's limb and the duskside terminator are overlaid on the picture.

turning-around of solar wind particles and their entry into the tail occurs, a series of powerful electromagnetic changes accelerates them to very high velocities down into the nightside ionosphere along the geomagnetic field lines. The particles which are precipitated into the auroral ionosphere include both electrons and ions.

The first detection of such particles was from sensors aboard rockets[17,18]. Since then, there has been phenomenal progress in the relevant instrumentation, and it has now become possible to distinguish between electrons and ions in terms of both their number flux and their energy. Fig. 6 shows the energy flux in Watts, and the number flux in electrons/sec-ster, within the entire auroral zone for different magnetic activity levels ranging from Kp = 0 to Kp = 6 [after Hardy *et al.*[19]]. Hardy *et al.*[20] showed a similar figure for ions (not depicted here). Various measurements suggest that the electron energy flux exceeds the

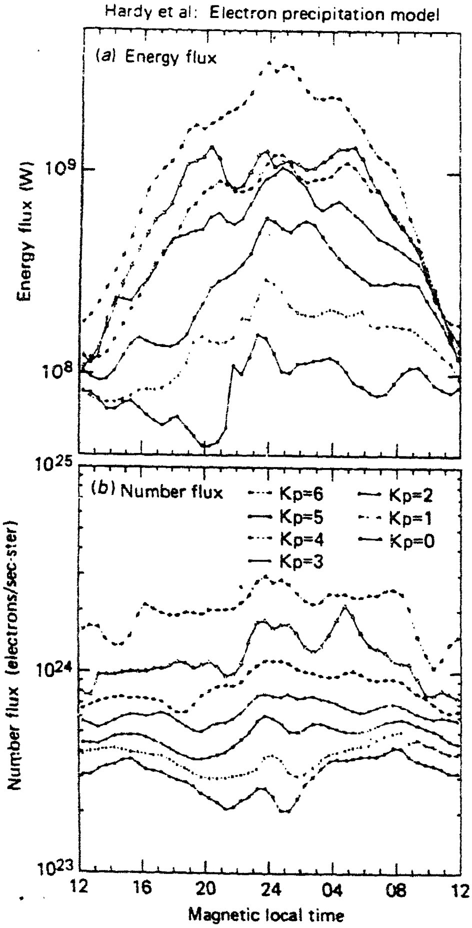


Fig 6 The energy flux and number flux variation with local magnetic time, for values of Kp varying from 0 to 6. While the number flux remains fairly constant in time, the energy shows a marked increase (more accentuated during magnetic disturbance) during magnetic midnight as compared to magnetic noon[19]

ion energy flux by a factor of less than 10, but the number flux of electrons exceeds that of ions by a factor[21] of almost 50. The features which cause acceleration of electrons and ions into the auroral ovals are discussed in the sections which follow.

The Solar Wind-IMF-Magnetosphere Electrical Generator

In the previous section, it was mentioned that the solar wind flows all around the obstacle presented by the geomagnetosphere. It must be kept in mind that this obstacle is a three dimensional structure, whose surface (the magnetopause) is constituted by the geomagnetic force-field. If the geomagnetic field was very weak like the magnetic field of Venus, and if the Earth had an ionosphere of low electron density like Venus, that would have been the end of the story. The solar wind would have merely flowed around the weak magnetosphere and past it, much like a stream of water round a spherical stone. As matters stand, the Earth has a substantial ionosphere (to be discussed in a later section), and a sizable magnetic field which interacts with the solar emissions—hence it has aurora. The recent VOYAGER 2 mission showed that the outer planets Jupiter, Saturn, Uranus and Neptune also have substantial magnetic fields and auroral activity[22,23].

At this stage we bring in the role of the Interplanetary Magnetic Field (IMF). The IMF emanates from the Sun's equator in spiral shape, and because of the alternating sector structure discussed under 'Solar Emission', it can be directed either southwards in the ecliptic plane or northwards. In this epoch, the Earth's magnetic field is directed northwards with respect to the ecliptic plane. If the IMF be also directed northwards, then there is no interaction between the two magnetic fields. The Earth's magnetic field lines then simply drape around the surface which is known as the Magnetopause and form a cavity which is isolated from the IMF. If the IMF is directed southwards however, it merges readily with the geomagnetic field, across the magnetopause boundary, as was shown in Fig. 4.

In Fig. 4, the thick arrows show the flow of the solar wind above and below the Magnetopause; solar wind also flows along the equatorial sides of the magnetopause (labelled as the Low Latitude Boundary Layer). It is clear that the solar wind flows across the IMF-geomagnetic field linkage, only when the IMF is directed southwards. As the solar wind consists of free electrons and protons, it behaves like a conducting element moving through a magnetic field, and this results in generation of a voltage. The solar wind flowing across the IMF-magnetosphere linkage basically constitutes a Magnetohydrodynamic (MHD) generator, in which a flow of highly ionised gas through a magnetic field generates a Voltage.

When EMF is developed in a generator, there must be two terminals—in the case of the solar wind-magnetosphere generator, the morning side of the low-latitude boundary acts as the positive terminal, while the evening side of the low-latitude boundary acts as the negative terminal. Tapping a current from these terminals requires wires which connect the terminals to a load or resistance. In the present case, the load is the auroral ionosphere i.e., the auroral region to which the last closed field lines from the distant magnetosphere map, (cf. Fig. 7). The connecting wires are the highly-conducting geomagnetic field lines which participate in this ionosphere-magnetosphere mapping. The currents which flow are called the Field Aligned Currents (FAC), and in order to complete the circuit, current has to flow into the ionosphere along one set of field lines, and out along the other set. This auroral discharge circuit has been depicted very clearly by Akasofu and Kamide[3] and is shown in Fig. 7.

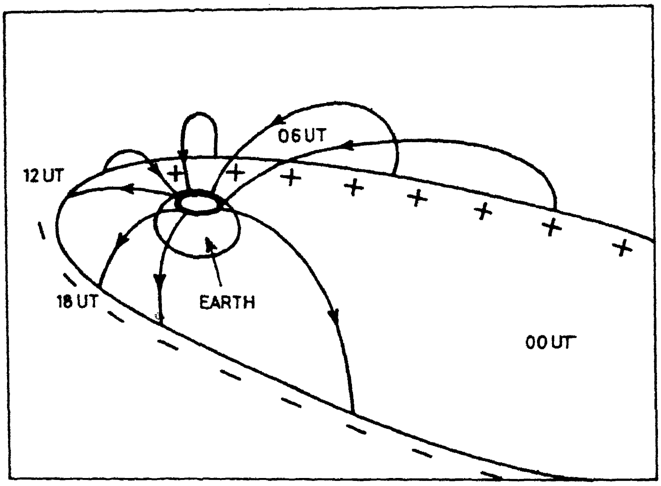
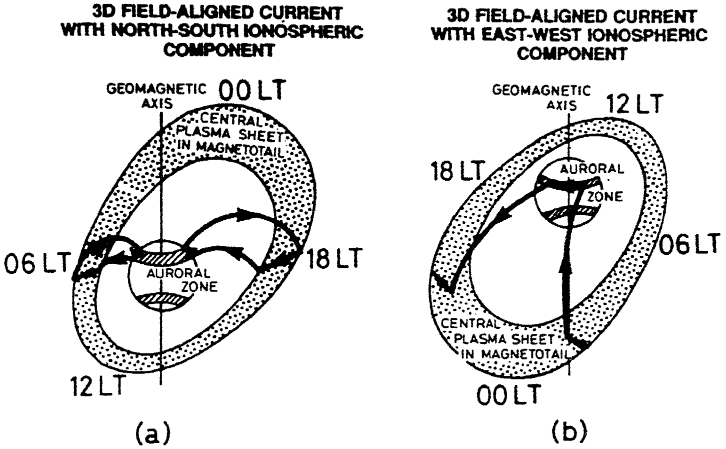


Fig 7 The solar wind IMF geomagnetic field power generator, causes accumulation of positive charge on the Dawnside and negative charge on the Duskside. The excess charge is discharged along geomagnetic field lines through the highly conducting auroral oval[3]

The first suggestion of field-aligned currents feeding into the auroral ionosphere came from Birkeland[24]. This was strongly supported by Alfvén[25], but Chapman[26] felt that the electric currents were confined to the Earth's immediate vicinity, namely the ionosphere. Satellites finally proved that substantial field-aligned currents did flow into the auroral ionosphere from deep space[27,28], and hence justified the thinking of Birkeland and Alfvén. A historical account of this is given by Potemra[29].

Studies over the past two decades by the TRIAD, MAGSAT and HILAT spacecraft indicate that there are predominantly three categories of field-aligned currents. Field-aligned currents at the poleward side of the auroral region are referred to as *Region* 1 currents and those at the equatorward side of the auroral oval as *Region* 2 currents. The third category comprises currents near the dayside *Polar cusp* but located poleward of the *Region* 1 currents. Field-aligned currents near the midnight region of the auroral oval called the *Harang discontinuity*, are very complicated too, and can be regarded as falling into a fourth category. These field-aligned currents show a clear local time and seasonal dependence[30], and can occur as single sheets (either towards or away from Earth) or as double sheets (both towards and away from Earth).

Clearly there have to be closure currents and these must flow in the auroral ionosphere with its enhanced electrical conductivity. In their simplest forms,



3D FIELD-ALIGNED CURRENT WITH NORTH-SOUTH IONOSPHERIC COMPONENT

3D FIELD-ALIGNED CURRENT WITH EAST-WEST IONOSPHERIC COMPONENT

Fig 8 The field-aligned currents which originate at the inner and outer edges of the Central Plasma Sheet flow into and out of the auroral zone. This ionospheric closure current in the auroral zone can be of the north-south type as shown in Fig. 8(*a*), or of east-west type as shown in Fig. 8(*b*).

two basic types of 3-dimensional current systems can be visualised as shown in Fig.8(*a*) and 8(*b*). In Fig. 8(*a*), the ionospheric closure current flows in the north-south direction; satellites and ground-based magnetometers generally notice southward flow in the pre-noon ionospheric sector and northward flow in the post-noon ionospheric sector. The currents are believed to originate at the inner and outer edges of the Central Plasmasheet shown in Fig. 8(*a*) and 8(*b*).

The second type of field-aligned current is shown in Fig. 8(*b*). Here the current which normally flows from dawn to dusk in the tail gets disrupted due possibly to a highly resistive region forming in the tail—this again is an area of research investigation. The disruption causes current to flow from the tail into the auroral ionosphere on the dawnside, then westward in the auroral zone as the electrojet, and finally back into the tail along the field lines from the duskside. This type of flow in the midnight sector often signifies the onset of what is called a *magnetospheric substorm*, namely a rapid transfer of electromagnetic energy from the magnetosphere to the ionosphere. This region of onset is referred to as the *Harang discontinuity*.

There is a third type of current flow in the auroral ionosphere which seems to consist of circular loops of current, originating mainly in the midnight sector and drifting towards the Dawnside and the Duskside local time sectors. Its origin lies in a combination of Pedersen and Hall conductivities, and we shall not deal with that aspect here. An up-to-date account of the 3-dimensional current systems at high latitudes is given by Iijima[31].

The Voltage generated across the Dawn-Dusk terminals by the solar wind-IMF-magnetosphere generator is of the order of 10^5 Volts, and the current which flows is about 10^6 Amperes. The power associated with this huge natural MHD generator is thus about 10^8 kilowatts—in comparison, the annual consumption of electric power in the USA is about 10^{12} kilowatt-hours. Finally it should be recapitulated that the solar wind-IMF-magnetosphere generator works with peak performance when the IMF is directed southwards. Northward IMF will merely result in the IMF field lines draping over the dayside magnetopause, and the geomagnetic field lines draping over the nightside magnetopause, with no reconnection between the two. Any power generated will then be within the magnetosphere, after the solar wind enters the nightside neutral sheet, and attempts to flow across the closed geomagnetic field lines.

Nightside Acceleration of Charged Particles into the Auroral Ionosphere

In the previous section it was mentioned that the field-aligned currents which link the ionosphere and magnetosphere; always have a segment in the auroral ionosphere, because it is highly conducting. It is necessary to understand why the ionosphere in this region should be so conducting, and for this we refer back to Fig. 4. A careful examination of this figure will show that magnetic field lines which start out from the equatorward and poleward borders of the auroral ovals, map to the region designated as the Central Plasma Sheet (CPS). This is the region of the magnetosphere adjacent to the Low Latitude Boundary Layer, and on the midnight side, lies close to the region where the solar wind is funnelled Earthwards through the tailside neutral sheet.

Irrespective of whether field-aligned magnetosphere-ionosphere currents flow or not in the Dawn and Dusk sectors, for a southward IMF a potential difference does exist between the Dawn and Dusk sectors. The electric field E associated with this is directed from Dawn to Dusk, and is at right angles to the geomagnetic field $\tilde{B}$ in the equatorial plane. The charged particles of the solar wind which have entered the nightside CPS are accelerated Earthwards with a high velocity $\tilde{V}$ given by

$$\tilde{V}=\tilde{E}\times\tilde{B}\,.$$

Once they are well inside the closed field line domain, the charged particles prefer to flow along the field lines rather than across them, and the entire CPS region gets filled with energetic charged particles. Their high velocities drive them towards the Earth along the field lines, but it is believed that a second accelerating region is required at 10,000-20,000 km above the auroral regions in order to given them high velocity. It is this potential which is believed to precipitate the charged particles into the auroral ionosphere with energy enough to 1) cause a great deal of secondary ionisation and render it highly conducting and 2) send the gaseous atoms and molecules of the auroral atmosphere into excited states, and thereby induce them to give off the optical emissions which constitute the visible aurora. The action of these field-aligned V-shaped auroral potentials is shown in Fig. 9 after Akasofu and Kamide[3].

The Onset of Auroral Activity

At first sight, auroral activity seems totally arbitrary, with aurorae of different shapes and hues occurring anywhere in the sky at anytime. Years of close study have shown however, that a definite pattern is followed, with some minor deviations.

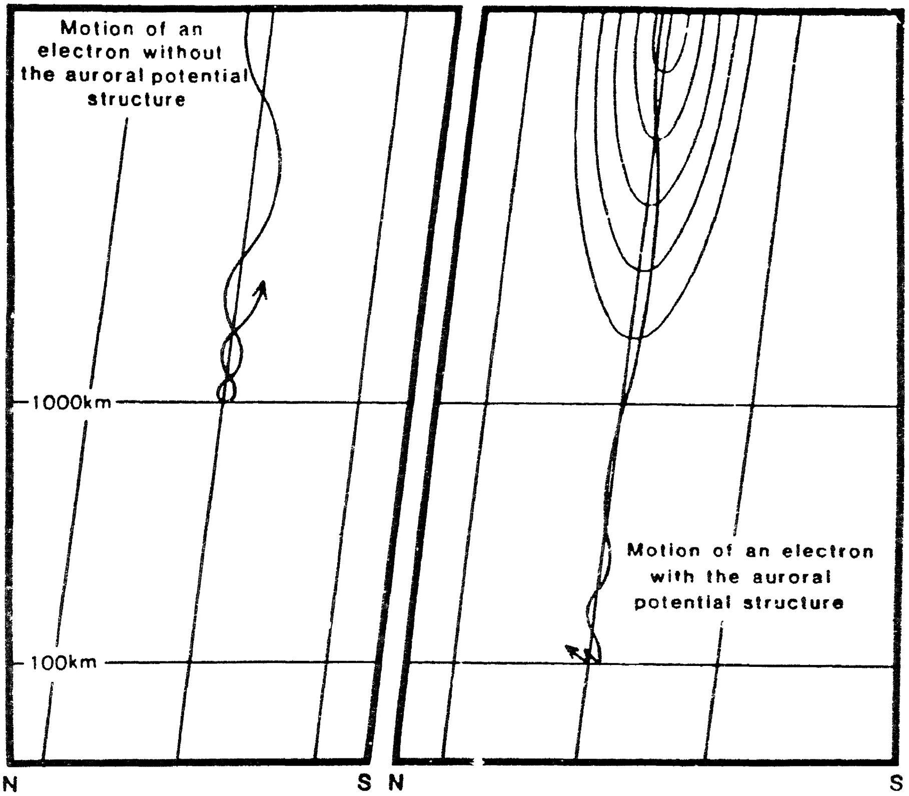

Motion of an electron without the auroral potential structure
1000km
100km
N
S
N
S
Motion of an electron with the auroral potential structure

Fig 9 While electrons are accelerated along the field lines in the nightside tail, it is believed that another accelerating region located at 10,000-20,000 km is required to bring them down to ionospheric altitudes. This region is believed to contain V shaped potential structures as illustrated in the figure[3]. The potential structures shoot the electrons straight down to ionospheric altitudes (100 km) instead of letting them mirror back at 1000 km altitude

For the reasons stated in Sec 6, the onset of auroral activity is always in the midnight sector, in a region referred to as the *Harang discontinuity.* The first sign is a brightening of an auroral structure with development of a ray and wave pattern. Within a few minutes the structure moves sharply northwards, and curls into a form with a bright ball-like head towards the west. This form moves rapidly westward towards the Dusk sector, and is called the *Westward Travelling Surge* (*WTS*). Some of the structures being formed in the midnight side also drift eastwards towards the Dawn sector, assuming the shape of inverted omegas, and hence are called *omega bands.* With time, the auroral structures in all local time sectors disintegrate into patchy luminescence, and that signifies the end of that episode of auroral activity—the whole is known as an auroral substorm, and its magnetic signature will be discussed in a later section. The entire sequence of events is repeated again if another substorm follows.

The auroral substorm is a clear indication that the IMF is directed in such a manner that the solar wind-IMF-magnetosphere generator is working very effectively, and is generating strong Dawn-to-Dusk electric fields. In other words the processes of flow of electromagnetic energy from 1) the Sun to the interplanetary medium, 2) from the interplanetary medium to the geo-magnetosphere, and 3) from the magnetosphere to the geo-ionosphere, are operating with great vigour.

The Formation of Optical Aurora in the Ionosphere

It was mentioned in an earlier section that the phenomenon of aurora can occur between 100-1000 km altitude, in the regions which are known as the Bottomside and the Topside ionosphere. Optical emissions are generated when the charged particles of solar wind origin, plummet into the ionosphere after 1) being accelerated near the outer edge of the CPS by the strong Dawn-to-Dusk electric field generated by the solar wind-magnetosphere generator and 2) being further accelerated along magnetic field lines by V-shaped potential structures located about 10,000-20,000 km above the auroral ionosphere. The accelerated electrons lose energy through collisions with the atoms and molecules of the O_2, N_2 and NO gases in the auroral atmosphere. Typically a 300 eV electron from outer space penetrates to altitudes of about 400 km, and electrons with energy 3-10 keV come down to about 100 km altitude.

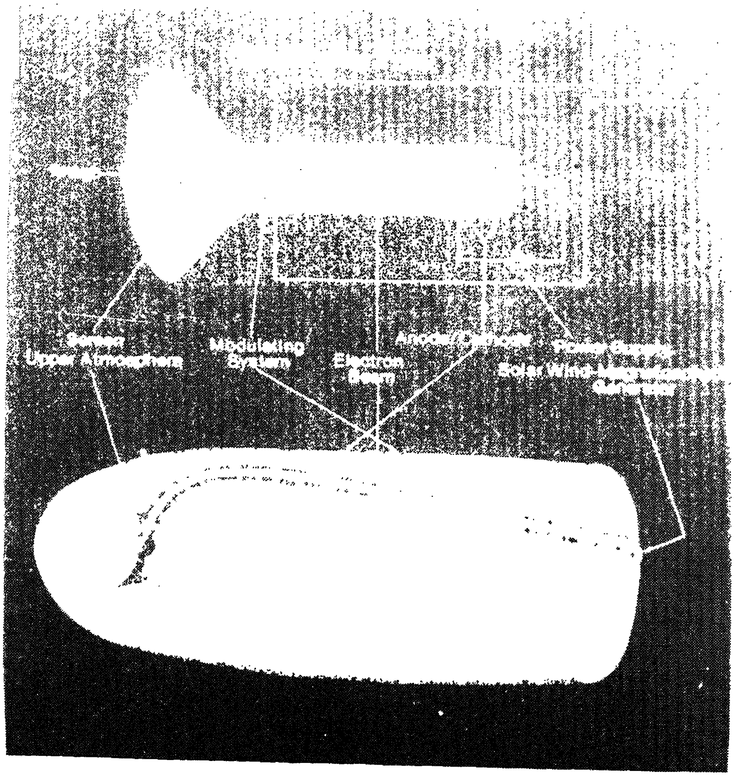


Fig 10 Since the 1960s it has been thought that the auroral ionosphere with its images of aurora, acts like the screen of a television tube. Today this is understood to be a reality, and a clear comparison of the various parts of the ionosphere-magnetosphere system with the components of a Cathode Ray Tube is shown in the figure[33].

Electrons can hit N_2, O_2, or NO molecules and atoms of nitrogen and oxygen. In each case they impart their energy to the gases and 1) ionise them, knocking out one or two electrons out of their orbits and 2) sending the ionised gases into "excited" states i.e., higher energy states. It is in returning from the "excited" state to the "normal" state that the ionised N_2 molecule emits strong violet-green light at 3914Å, the ionised O atom emits greenish-white light at 5577Å. At lower heights of 90 km, more energetic electrons (>10 keV) cause N_2 molecules to emit pink or crimson-red light.

A typical example of the auroral emission process is

$$N_2 + e\,(\text{energy of few keV}) \rightarrow N_2^+\,(\text{excited}) + e + e\,(\text{with energy still a few keV})$$

$$N_2^+\,(\text{excited}) \rightarrow N_2^+\,(\text{ground state}) + \text{Optical emission}\,(3914\text{Å})$$

A summary of the various emissions which are possible from auroral processes is given by Vallance-Jones[32].

The Auroral Magnetosphere-Ionosphere System as a Cathode Ray Tube

Fig. 10 from Akasofu[33] is a very clear depiction of the analogy between a Cathode Ray Tube (CRT) and the field-line linked Ionosphere-magnetosphere system, and is almost self-explanatory. The Solar wind-Magnetosphere Power generator discussed under 'Solar wind-IMF-Magnetosphere' serves as the Power Supply with its terminals in the Dawn and Dusk sides of the Earth's equatorial plane (shown in Fig. 10). The Anode-Cathode system in the CRT which accelerates electrons to high velocity is believed to be located at 1-2 R_e away from Earth, and is referred to under 'Nightside Acceleration of Charged Particles' as the *V*-shaped Potential structure. The Modulating System in the CRT which decides just how much the electron beam is bent could be located anywhere along the connecting geomagnetic field lines and its location has not been understood as yet. It is logical to believe that the degree of modulation (i.e., bending of the electron beam so as to cause auroral structures of different shapes) would be decided by the strength of electric and magnetic fields in the geomagnetosphere, and hence by solar field and particle emissions. In the case of the CRT, the screen is coated with fluorescent material which glows when it is hit by charged particles. In the real Ionosphere-Magnetosphere system, the screen is the upper atmosphere with its gases of O_2, N_2 and NO in atomic and molecular form, which "fluoresce" through ionisation and excitation.

This analogue between the CRT and the auroral Ionosphere-Magnetosphere system is a beautiful one, and it makes us realise that planet Earth lies within a powerful electromagnetic generator, and is basically cocooned by its gaseous atmosphere.

Geomagnetic Records as a Pointer to the Electromagnetic State of the Ionosphere-Magnetosphere System

Apart from optical footprints in the auroral ionosphere, the electromagnetic changes in the magnetosphere brought about by the solar-terrestrial energy exchange process have other manifestations e.g., as electric fields, magnetic fields, X-ray and VLF emissions, strong absorption in radio waves etc. Here we would like to focus on the importance of geomagnetic records (magnetograms) which can be ground-based or satellite-based, in understanding the electromagnetic state of the geospace environment.

A magnetometer measures the magnetic effects of the integrated currents around it, and is most responsive to currents in the nearest proximity. In the case of a ground-based magnetometer in the auroral region, the location of these currents is in the ionosphere at about 100-110 km altitude. Field-aligned currents which flow into the auroral ionosphere from the deep magnetosphere are also capable of leaving magnetic signatures on ground. A vertical current leaves signatures in the *Y* (or *D*) and *X* components, a north-south current leaves variations in the *Y* (or *D*) and *Z* components, while an east-west current leaves variations in the *X* and *Z* components. Thus magnetometers by themselves can never uniquely define a current system, and hence they need to be supported by other instruments such as *a*) Riometers which detect radio wave absorption and hence detect field-aligned currents at auroral latitudes, or *b*) Electric field detectors. If any two of these are used simultaneously, it is possible to obtain a better understanding of the 3-dimensional current systems operating in the auroral ionosphere.

Another important feature which an array of magnetometers at auroral locations can detect is the velocity of mobile current systems which sweep overhead as auroral structures. Thus if there are 3 magnetometers operating at the vertices A, B, C of a triangle, and if an auroral current sweeps through the ionosphere above, passing first over A, then B, then C, then it is clear that there will be certain time-lags in the magnetic signatures left at A, B and C. If A, B and C be rotated from the geographic into the geomagnetic coordinate system which the currents follow, then knowing the time-lags in the signature, and the exact distances between A, B and C, one can determine the velocity (i.e., speed and direction) of the auroral current which swept overhead. India is carrying out just such an experiment at Maitri and DG, the Indian Antarctic stations, and a third location which varies from year to year. Shown in Fig. 11[34] is an example of the time-delay in the Ps6 pulsations recorded at seven Canadian auroral stations located in the magnetic north-south and east-west directions. The velocity of the overhead auroral current system obtained in this case was ~ 1 km/sec.

The velocity obtained in the ionosphere can be extrapolated into the distant magnetosphere by using mapping factors which define the projection[34]. This is because the geomagnetic field lines link the ionosphere and magnetosphere, and the mapping factor is decided by the geometry of the field line. The mapping factor between the midnight auroral region and the nightside magnetospheric tail, is almost 3200. This is the factor by which velocities and scale-lengths observed in the ionosphere are amplified in the magnetosphere, and it is thus that we use auroral ionospheric signatures to estimate physical processes in the deep distant magnetosphere.

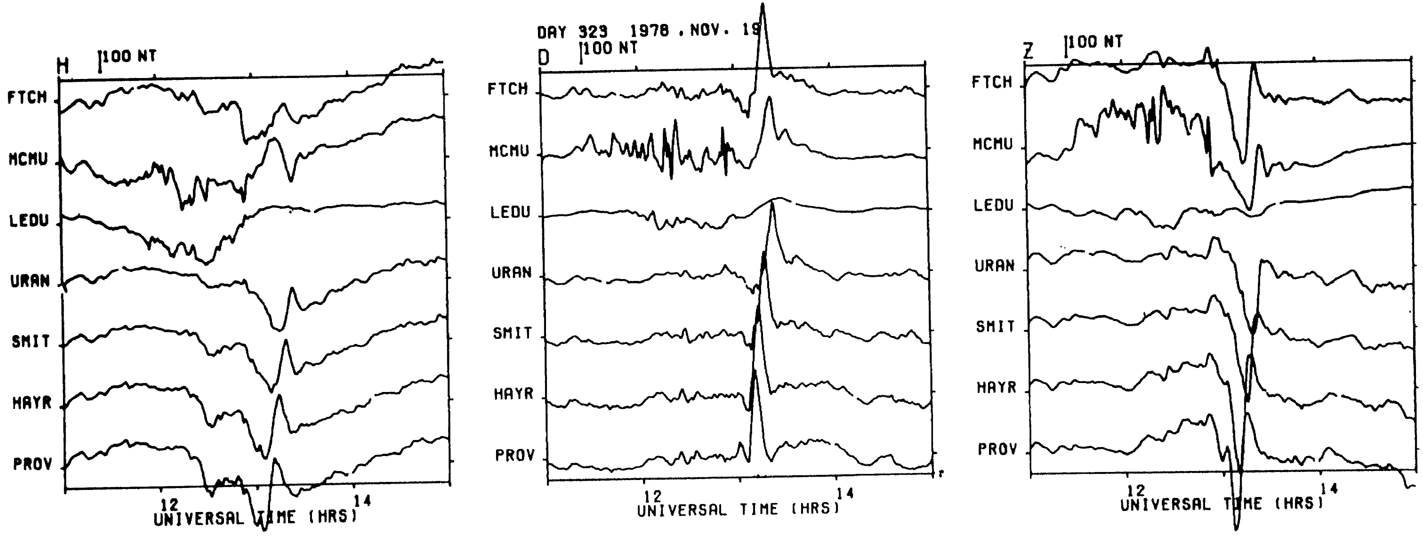


Fig 11 Timelags can be clearly seen in all three components *H*, *D*, *Z* of a pulsation which occurred after 13:10 UT, on 19 Nov 1978, at the lower four locations depicted in the figure. The 7 stations formed the Alberta array of magnetometers; the lower four stations URAN, SMIT, HAYR and PROV covered 12° longitude in the east-west direction, and caught an electromagnetic disturbance propagating eastwards[14]

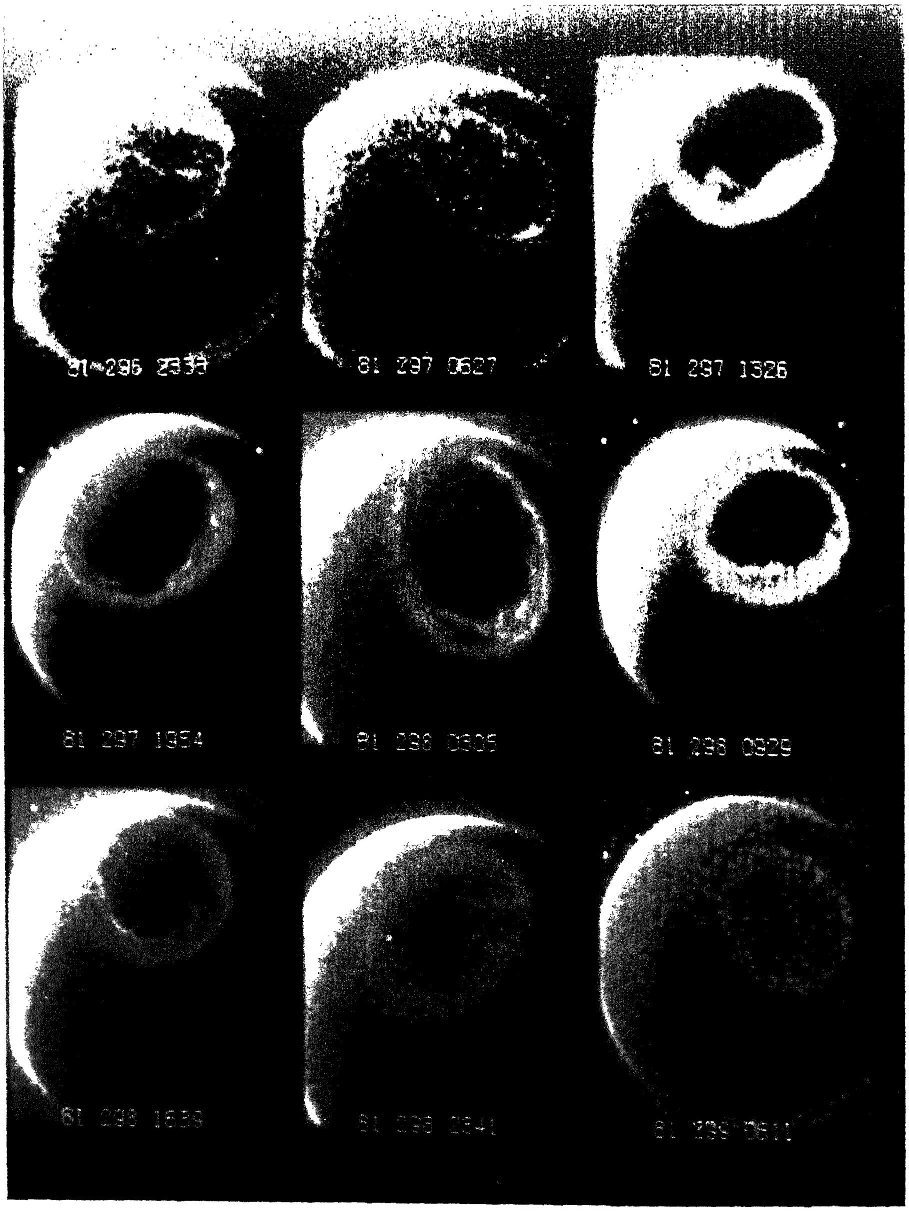


Fig 12 The Dynamics Explorer (DE) satellite captured this view of the entire auroral oval during nine consecutive orbits. Below each frame is the year, the day number and the time in UT. The first two and last two frames correspond to northward interplanetary magnetic field (IMF), and exhibit theta type aurora located poleward of the normal auroral oval. The other seven frames were taken during conditions of southward IMF, and show active auroral conditions with an auroral substorm in progress[16]

Conclusions

The first time that the auroral ionosphere was compared to a TV tube seems to have been in the 1960s[4]. Observations at that time suggested that polewards of the 70° magnetic latitude, auroral structures drifted anti-sunwards, while equatorwards of this latitude they seemed to drift sunwards. This drift pattern was interpreted as mimicking of the plasma convection pattern in the deep magnetosphere, where in the low latitude boundary layer, the solar wind moved towards the magnetospheric tail, and within the Central Plasma Sheet, solar wind moved towards the Earth.

Today the International Space Science community has come a long way from the early observations of Hiorter and Celsius of 1 March 1741, of a freely suspended magnetic needle oscillating violently, in concert with the variations in brightness of an auroral display in progress at that time. The suggestions of Gauss in 1883, in his "Allgemeine Theorie des Erdmagnetismus" of "electric currents, which are sent out from the Sun into space", and of Balfour Stewart who in 1882 suggested that the daily variations in the Earth's magnetic field could be due to currents which flowed in the ionised atmosphere at high altitudes, have since been vindicated many times over by ground and satellite-based observations. Kristian Birkeland's suggestions[24] in 1908 of field-aligned currents flowing into high-latitudes have been repeatedly confirmed by the TRIAD, MAGSAT and HILAT spacecraft (hence the name Birkeland currents for the 3-dimensional auroral current systems).

The analogy between a TV tube and the magnetosphere-ionosphere system was proved correct by global views of the entire auroral oval provided by satellite imagery. Time-sequential views of auroral lights flitting about on the "auroral ionospheric" screen have been captured spectacularly by the DMSP, HILAT, Polar BEAR, DE1 and VIKING satellites (Fig. 12), and future endeavours in this aspect are covered by Williams[35]. Two space experiments executed recently in March 1992 on the ATLAS (Atmospheric Laboratory for Applications and Science) mission were electron beam injections for the creation of artificial aurora[36] and imagery of the artificial aurora thus created. The 50 Colour Plates of satellite images of the aurora in Meng *et al.*[16] illustrate the grandeur and beauty of the phenomena, and the reader would do well to go through these.

References

1 S I Akasofu and D S Kimball *J atmos terr Phys* **26** (1964) 205-211

2 J A Ratcliffe In: *Sun, Earth and Radio—An Introduction to the Ionosphere and the Magnetosphere* World Univ Library Weidenfeld and Nicolson 5 Winsley St London W1 (1970) 28-29

3 S I Akasofu and Y Kamide In: *The Solar Wind and the Earth* (Eds S I Akasofu and Y Kamide) *Geophysics and Astrophysics* Publ Terra Scientific Pub Co Tokyo (1987) 143-159

4 G L Siscoe In: *Auroral Physics* (Ed C I Meg *et al.*) (1991) 159-175

5 C W Snyder, M Neugebauer and U R Rao *J geophys Res* **68** (1963) 6361-6370

6 J W Kern In: *Physics of Geomagnetic Phenomena* Vol 2 Academic Press N Y London (1967) 1037-1109

7 P J Coleman Jr, C P Sonett, D L Judge and E J Smith *J geophys Res* **65** (1960) 1956-1857

8 J P Heppner, N F Ness, T L Skillman and C S Scearce *J phys Soc Japan* **17** (1962) Suppl A-II, 546-559

9 H Bridge, C Dilworth, A J Lazarus, E F Lyon, B Rossi and F Scherb *J phys Soc Japan* **17** (1962) A-II 553-560

10 J L Burch In: *The Solar Wind and the Earth* (Eds S I Akasofu and Y Kamide) Terra Scientific Publ cc. Tokyo (1987) 103-122

11 R T Merrill and M W McElhinny In: *The Earth's Magnetic Field—Its History, Origin and Planetary Perspective* Academic Press London N Y (1983) ch 12

12 N F Ness and J M Wilcox *Science* **148** (1965) 1592

13 D H Fairfield *Space Res* **8** (1968) 107

14 P J Kellogg *J geophys Res* **67** (1962) 3805-3811

15 W I Axford *J geophys Res* **67** (1962) 3791-3796

16 C I Meng, M J Rycroft and L A Frank (Eds) *Auroral Physics* Cambridge Univ Press (1991)

17 C E McIlwain *J geophys Res* **65** (1960) 2727-2747

18 L R Davis, O E Berg and L H Meredith In: *Space Res* 1 (*Proc COSPAR Space Sci Symp*) North Holland Publ Co Amsterdam (1960) 721-735

19 D A Hardy, M S Gussenhoven and E Holeman *J geophys Res* **90** (1985) 4229-4248

20 D A Hardy, M S Gussenhoven and D Brautigam *J geophys Res* **94** (1989) 370-392

21 P T Newell, C I Meng and D A Hardy In: *Auroral Physics* Combridge Univ Press Cambridge, N Y Melbourne (1991) 85-95

22 J T Bergstralh *Rev Geophys* **25** (1987) 251-259

23 O Lindley *Nature (London)* **341** (1989)

24 K Birkeland *The Norwegian Aurora Polaris Expedition* 1902-1903 Aschehoug Oslo Norway (1908)

25 H Alfvén *Theory of Magnetic storms Kungl Sv Vet Akademiens Hand* **18** (1939) No 37

26 S Chapman *Terr Magn atmos Electr* **40** (1935) 349

27 A J Zmuda, J H Martin and F T Heuring *J geophys Res* **71** (1966) 5033

28 A J Zmuda, F T Heuring and J H Martin *J geophys Res* **72** (1967) 1115

29 T A Potemra *Magnetospheric Currents, John Hopkins APL Tech Dig* **4** (1983) No 4 Maryland USA

30 R Fujii, T Iijima, T A Potemra and M Sugiura *Geophys Res Lett* **8** (1981) pp 1103-1106

31 T Iijima In: *Auroral Physics* (Ed C-I Meng *et al.*) Cambridge Univ Press Cambridge N Y Melbourne (1991) 401-408

32 A Vallance-Jones In: *Auroral Physics* Ed C I Meng *et al.*) Cambridge Univ Press Cambridge N Y Melbourne (1991)

33 S I Akasofu *Aurora Borealis-the Amazing Northern Lights, Alaska Geogr Soc* **6** (1979) No 2

34 G Rajaram, G Rostoker and J C Samson *Planet Space Sci* **34** (1986) 319-329

35 D J Williams In: Auroral Physics (Ed C I Meng *et al.*) Cambridge Univ Press Cambridge N Y Melbourne (1991) 449-456

36 J L Burch, S B Mende, N Kawashima, W T Roberts, W W L Taylor, T Neubert, W C Gibson, J A Marshal and G R Swenson *Geophys Res Lett* **20** (1993) 491-494

37 S B Mende, J L Burch, G R Swenson, E K Aamodt, S P Geller, R L Rairden and P L Hassler *Geophys Res Lett* **20** (1993) 495-498

PHYSICS AND TECHNIQUES OF REMOTE SENSING FOR WEATHER AND CLIMATE RESEARCH

P C PANDEY
Space Applications Centre, Ahmedabad-380053, *India*

(*Received* 02 *July* 1993; *Accepted* 05 *July* 1993)

Space based remote sensing techniques during the last three decades have made important contributions to the understanding of our planet earth. Visible, infrared and microwave regions have been used extensively for the remote sensing of the atmospheric and oceanographic parameters needed for weather and climate research. Microwave sensors due to their ability to propagate through the atmosphere without being influenced very much have drawn the attention of oceanographic community in particular and atmospheric scientists in general for providing quantitative information. This paper deals with research primarily carried out by Indian scientists using microwaves, both passive and active, to study our planet earth together with satellite system planned for future and various international programs whose data would be useful for studies of monsoon.

Key Words: **Remote Sensing; Weather; Climate; Visible; Infrared and Microwave Regions**

Introduction

Three decades of satellite observations of the earth have provided much information of the different phenomena and over the years the need to understand the earth as a system has developed. NASA has already planned 'Mission to Planet Earth' and space agencies of Japan and Europe have joined them in the overall effort of monitoring the earth as a system. In India, the launch of INSAT series of satellites for weather monitoring, IRS series for earth resources applications and some of the earlier satellites such as Bhaskara for oceanographic applications, have changed the perspectives of both the scientists and the informed public regarding weather prediction and natural resources management.

Mathematical models are the basic tools for prediction which can simulate, describe and predict natural processes. These models require input data and time series of global observations from space with selected simultaneous *in situ* observations are indispensable for the models to become reliable. Observations are also needed to simplify and parametrise the complex physical processes so that they can be incorporated into models. Observations over the atmosphere and oceans are important. The observations over the oceans are very sparse and there appears to be no immediate prospects of improving it by conventional *in situ* observations. Under the above circumstances, the data from satellites are all the more necessary. For the study of regional phenomena such as monsoon in the Indian context, whose lucid description is given by Das, P K[1] and whose origin is believed to be in the oceans, we must monitor weather parameters over the oceans. The spectral peak of the incident solar radiation arriving at the earth is in the visible part of the spectrum, where atmosphere is relatively transparent. So over the major oceanic portion of the globe that part of the energy (~ 70%) that is not reflected directly back to space, is absorbed not by the atmosphere but by the upper few ten of meters of the sea. The major currents redistribute this amount of heat around the oceans, and subsequent release of heat to the atmosphere drives the atmospheric circulation. Ocean's role in climate studies is mainly represented by the large exchange of energy at the interface through air-sea fluxes, heat absorption and redistribution by ocean currents and gyres. Coupled ocean-atmosphere models are under development for global and regional scale studies.

Microwave sensors have the unique advantage of providing all weather data over oceans. It is always advantageous to use data from sensors operating at different wavelengths so that optimum advantage can be taken from each sensor. Certain ocean parameters such as wave height and ocean wave spectrum are unique to microwaves. Data obtained from our own satellites and satellites from the other countries have given us the opportunities to learn the physical basis of measurements and the retrieval algorithms to enable us to compare the geophysical variables obtained from satellites and *in situ* observations. In this review article, some examples will be given about the scientific results obtained using primarily passive and active microwave sensors. These experiences which have been gained will be useful to plan our own ocean/atmospheric earth observing mission and collaborate with scientists in other countries.

Physical Basis of Passive Microwave Radiometry

Microwave radiometry is the detection of natural thermal radiation. The propagation of electromagnetic radiation through the earth's atmosphere where it encounters absorption, scattering and thermal emission, is described by radiative transfer[1a]. The upwelling radiance in a monochromatic spectral band at the top of the atmosphere, can be written as

$$I_\nu(p=0)=\varepsilon_\nu B_\nu(T_s)\,\tau_\nu(p_s\rightarrow 0)+\int^{0} B_\nu[T(p)]$$

$$\times\frac{\partial\tau_\nu(p\rightarrow 0)}{\partial p}\,\mathrm{d}p+(1-\varepsilon_\nu)\,\tau_\nu(p_s\rightarrow 0)$$

$$\times\int_0^{p_s} B_\nu[T(p)]\,\frac{\partial\tau_\nu(p\rightarrow p_s)}{\partial p}\,\mathrm{d}p, \qquad \ldots(1)$$

where $\tau_\nu(p_i\rightarrow p_j)$ denote the transmittance functions. The pointing direction of the arrow indicates the direction of the atmospheric emission (e.g. $\tau(p_s\rightarrow 0)$ means transmittance from surface to space).

$\varepsilon_\nu B_\nu(T_s)$ = upwelling radiance at the surface
T_s = surface temperature
ε_ν = emissivity of the surface

$$\frac{\partial\tau}{\partial p}(p_i \to p_j) = \text{weighting function}$$

Eq. (1) is the fundamental equation describing the upwelling radiation at the top of the atmosphere. The first term denotes the surface contribution. Term 2 represents the direct atmospheric emission. As the atmosphere (to a first approximation) is radiating isotropically, it is not the only direct emission that contributes to the signal at the top of the atmosphere, that part of the emission which is reflected back by the earth's surface at the same frequency (term 3) has also to be taken account.

Obviously the observed radiance reflects the gaseous profiles, temperature and surface effects, so in principle it should be possible to derive the latter by inverting eq. (1). Eq. (1) looks slightly different in IR and microwave parts of the spectrum. In IR, the surface emissivity is taken as unity, therefore term 3 disappears in eq. (1). In the microwave region (1mm to 1m), the surface emissivity varies over a wide range, 0.4 to 1.0. Consequently term 3 cannot be neglected in this region. However, in the microwave region, Planck's function is linearly related to temperature and this is referred to as Rayleigh Jeans (RJ) approximation. Because of this reason, we use temperature for radiative transfer eq. (1). Using RJ approximation for a blackbody gives

$$B_\nu(T) = \frac{2kT}{} \qquad (2)$$

The RJ approximation for a blackbody. However, radiation from a gray body, for a given wavelength, is equivalent to radiation from a blackbody at lower temperature. This temperature is referred to as the brightness temperature T_B of the gray body. Thus

$$T_B = \varepsilon T \qquad \ldots (3)$$

Consequently eq. (1) becomes under RJ approximation as

$$T_B(0) = \varepsilon_\nu T_s \tau_\nu(p_s \to 0) + \int_{p_s}^{0} T(p)\frac{\partial\tau_\nu}{\partial p}(p \to 0)\,\mathrm{d}p + (1-\varepsilon_\nu)\tau_\nu(p_s \to 0)\int_{0}^{p_s} T(p) \times \frac{\partial\tau_\nu(p \to p_s)}{\partial p}\,\mathrm{d}p \qquad \ldots (4)$$

The emission from ocean surface under different wind conditions and for frequencies below 40 GHz, generally used for surface remote sensing is given by Pandey and Kakar[2]. The absorption, emission and transmission properties of the atmosphere are mainly due to the following effects.

—oxygen molecule
—water vapour molecules
—absorption and scattering by water droplets

The frequency dependence of these effects are known to useful extents[3] and microwave radiometry has been used for quantitative measurements of water vapour content and liquid water contents[4,5]. Oxygen, a major constituent of the atmosphere has a relative volume abundance of 0.21 whereas CO_2 which is a minor constituent has a relative volume abundance of 0.003. The measurements of microwave radiation at different frequencies with different oxygen absorption can be used for atmospheric temperature retrievals as is done with MSU instrument of the TIROS Operational Vertical Sounder (TOVS)[6,7]. Fig. 1 shows the microwave spectrum and the absorption lines selected for planned Advanced Moisture Sounder (AMSU).

Parameter Retrieval

The general scheme for parameter retrieval is shown in Fig. 2. The models of the atmospheric molecules, water vapour, oxygen and clouds, models of the ocean and ice are used to calculate emissivity, absorptivity and transmissivity which are then integrated by radiative transfer model to produce brightness temperatures, for a given frequency, polarisation and incidence angle. If antenna pattern effects are negligible, then brightness temperature approximates the antenna temperature. The most used retrieval methods are the one termed as Statistical retrieval methods and have been used for retrieving several meteorological and oceanographic parameters. The subject of retrievals has received considerable attention in recent years due to requirements of quantitative information for numerical models[8].

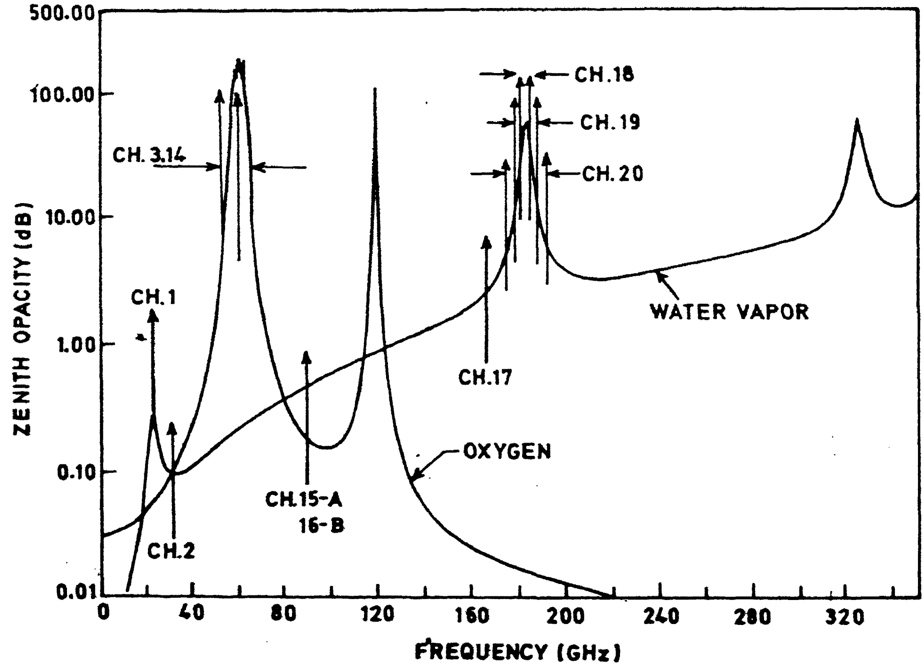


Fig 1 Atmospheric absorption lines and AMSU channels

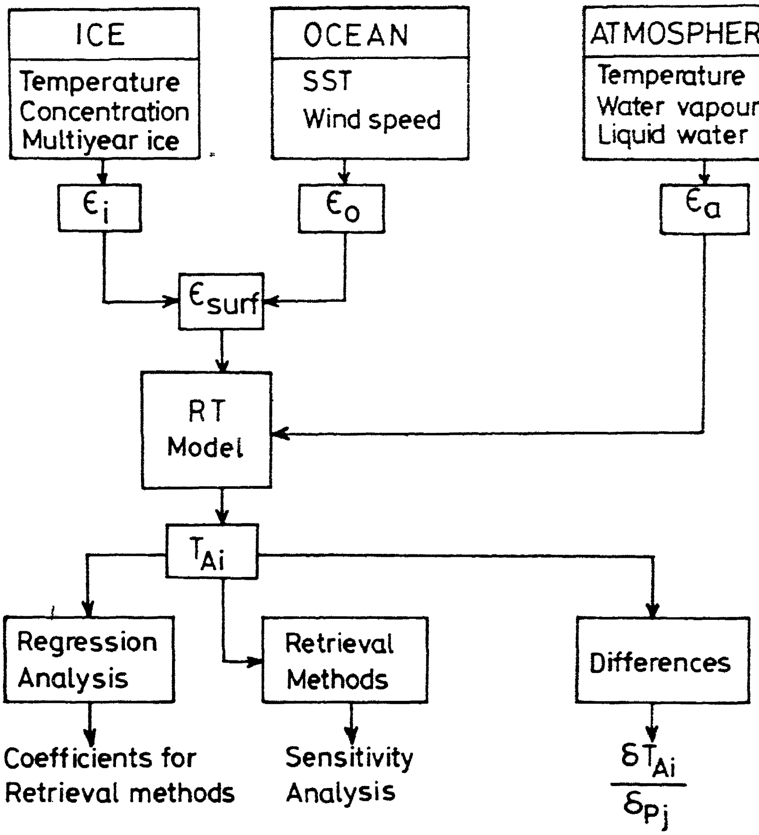


Fig 2 Context diagram showing the major components of radiative transfer computation and parameter retrieval

Some Examples of Scientific Results

Precipitable Water

The total atmospheric precipitable water over the oceanic regions has been studied extensively using passive microwave radiometer data. The satellite Microwave Radiometers (SAMIR) on board the Indian Satellites Bhaskara-I and II, have been used to derive precipitable water over the seas around India and its latitudinal variations[4,9]. The study was further extended by Gairola *et al.*[5] to compare SAMIR products with NOAA derived precipitable water. The study demonstrated the usefulness of all weather availability of moisture data over oceans and its application to map moisture field in a developing cyclone. The retrieval accuracy of total precipitable water with Bhaskara was comparable to that obtained from the previous similar missions[10].

The launch of SEASAT and NIMBUS-7 satellites by NASA with ten channels radiometer operating at 6.6, 10, 18, 21 and 37GHz with dual polarisations, have given global observations. The validation of precipitable water with *in situ* observations have given an rms error of 0.2-0.4gm/cm^2. A new algorithm was developed to retrieve various geophysical parameters[11] from multichannel observations. Fig. 3 gives the validation curve, and Fig. 4 gives the global distribution of moisture field[12] obtained using SEASAT-SMMR (Scanning Multi Channel Microwave Radiometer data). Cloud liquid water contents were also retrieved from these measurements, but they were not validated due to non-availability of frequent *in situ* measurements as in the case of humidity profiles.

Latent Heat Flux Study over Oceans

The total precipitable water derived from SEASAT-SMMR and the recent satellite DMSP (Defense Meteorological Satellite Program) of U.S.A. with SSM/I (Special Sensor Microwave Imager) giving the total precipitable water has been used by Gairola *et al.*[13,14], Gautam *et al.*[15], 1993 and Nimmi *et al.*[16] for the study of latent heat flux over Indian oceans. They modeled the surface level air humidity, a parameter required for flux calculations using Bulk Aerodynamic formultions, in terms of satellite derived product derived product of

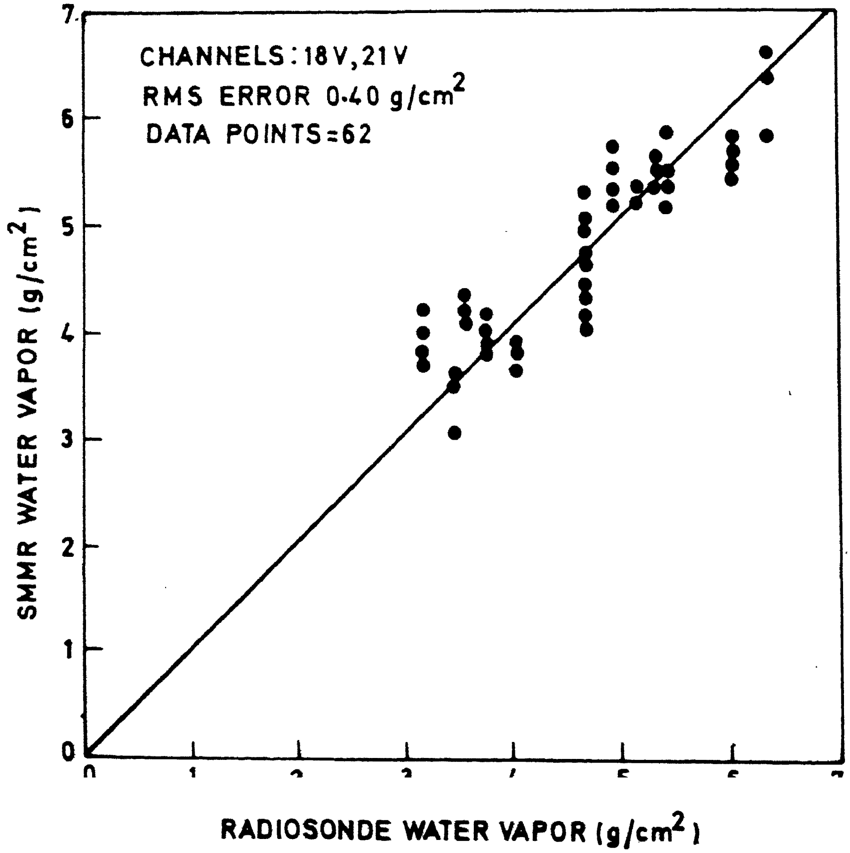


Fig 3 Comparison of precipitable water derived from SEASAT SMMR and nearby radiosonde for the month of September 1978 (after Pandey 1992)[12]

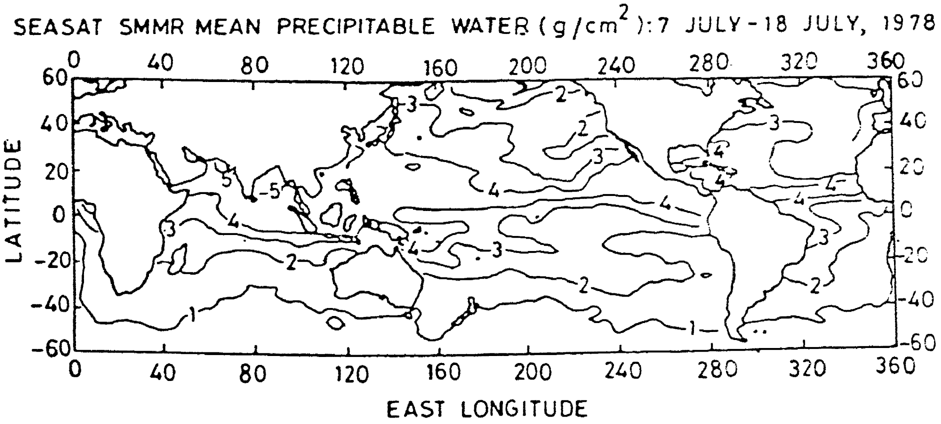


Fig 4 Global SMMR moisture field for the period July 7-July 18, 1978 (after Pandey 1992)[12]

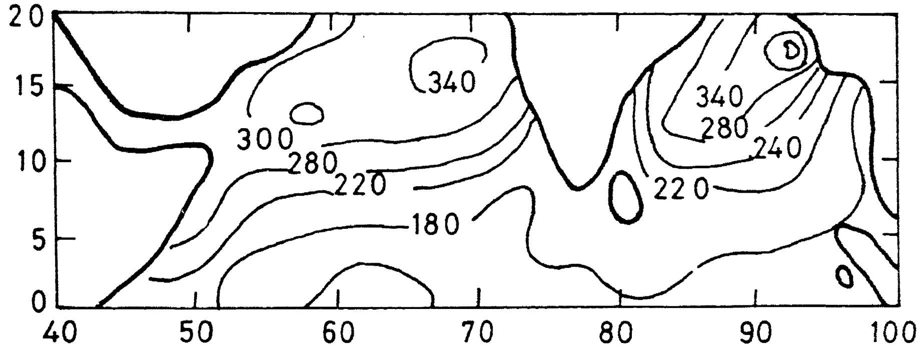


Fig 5 Latent heat flux (W/m^2) distribution for January 1988 from SSM/I and NOAA-AVHRR data over Indian oceans, (after Nimmi Nair *et al.* 1993)[16]

precipitable water. Liu *et al.*[17] had already demonstrated this type of approach for global computation of latent heat fluxes. These studies demonstrated the usefulness of passive microwave radiometry for air sea interaction study along with other satellite derived products such as SST from NOAA-AVHRR, and wind speed from either a scatterometer or passive radiometer. Fig. 5 gives a typical distribution of latent heat fluxes over Indian ocean. The accuracy of satellite derived fluxes with direct flux calculations using ships' observations gave an rms error of ~60 Watts/cm^2 using instantaneous observations. On monthly scales the accuracy of ~30 Watts/m^2 have been reported[18]. Simon and Joshi[19] studied air-sea interaction associated with a tropical cyclone using NOAA and GEOSAT data. In the absence of any direct observations over the oceans for such types of studies, satellites are going to play major roles in the future.

Kishtwal *et al.*[20] have also used satellite information to study surface level moisture transport during southwest monsoon. They have used NOAA/HIRS data for their study. Their analysis indicated a net positive surface level moisture flux divergence and negative surface level moisture flux divergence over the Bay of Bengal. Fig. 6 gives an example from Kishtwal *et al.*[20] Water vapour periodicities over Indian ocean during contrasting monsoons were also studied using NOAA satellite data by Kishtwal *et al.*[21] Efforts are underway to use satellite derived precipitable water alongwith characterization of the atmosphere from the past historical data to obtain humidity profile over oceans which forms important inputs to Numerical Weather Prediction Models.

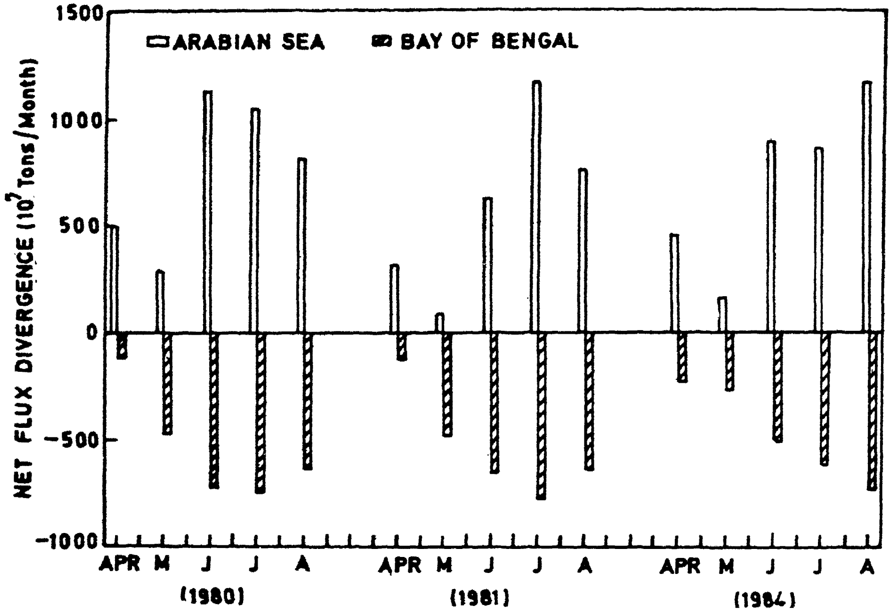


Fig 6 Flux divergence over the Arabian Sea and Bay of Bengal using NOAA HIRS data. (after Kishtwal *et al.* 1993)[21]

Wind Speed

Oceanic wind speeds are obtained from passive microwave sensors, scatterometers and radar altimeters. In case of passive radiometry, wind speeds are derived using multichannel observations since emissivity of ocean surface are wind speed dependent, in addition to wavelength, polarisation and observation angle. Fig. 7 shows the comparison of wind speed derived from SEASAT-SMMR with ships observations over North Atlantic during a special campaign, JASIN (Joint Air-Sea Interaction). The comparison gave an rms accuracy of 1.5m/sec[22].

SEASAT Scatterometer which was an active sensor operating at 14.6 GHz provided backscattering measurements at HH polarisation, which has been used to retrieve oceanic wind vector-speed and direction. Bragg scattering mechanism has been used to explain the radar return. The details of the instruments are described by Brown *et al.*[23]

Since the dependence of σhr° on ocean parameters are not fully understood, an empirical model function approach has been used to define the dependence

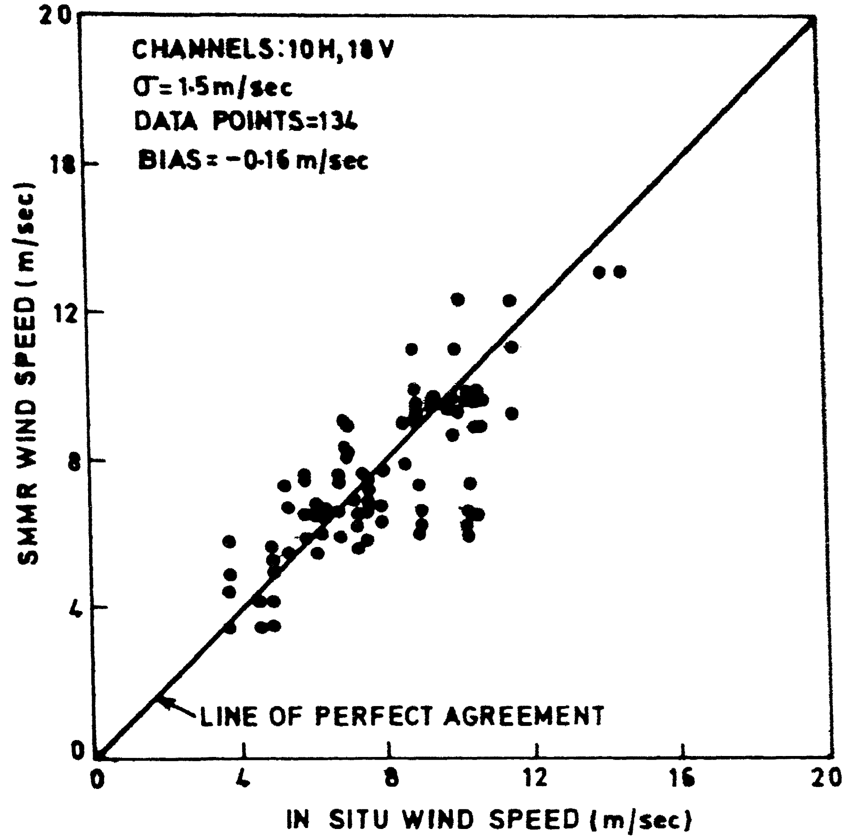


Fig 7 Comparison of wind speed derived from SEASAT SMMR data and its comparison with *in situ* observations during JASIN Experiments (after Pandey 1987)[22]

of σ° on windspeed incidence angle and azimuth angle Due to harmonic nature of dependence, multiple solutions are obtained which satisfy the measurements, posing a challenge to determine the unique solution. Gohil and Pandey[24] gave an algorithm for wind vector retrieval from two beam configuration such as SEASAT-SASS. The problem of ambiguity removal was minimised by making three measurements for a pixel such as in ERS-1 scatterometer operating at 5.3 GHz. Gohil[25] gave a median filter approach to remove the wind ambiguity and did simulations using three beam ERS-1 scatterometer configuration. He used the model function provided by ESA and which is different than the model function generated for SEASAT. Fig. 8 describes some simulation results from Gohil[25]. Sarkar and Raj Kumar[26] gave a new semiempirical sea spectrum for estimating the scattering coefficients which was later related with ocean surface wind.

Radar altimeter is another quantitative instrument to provide wind speed along a narrow swath below the satellite. Due to limited coverage, these measurements are useful for climatic studies. Gairola and Pandey[27] investigated the effect of foam and white caps on extraction of wind speed from pulse limited

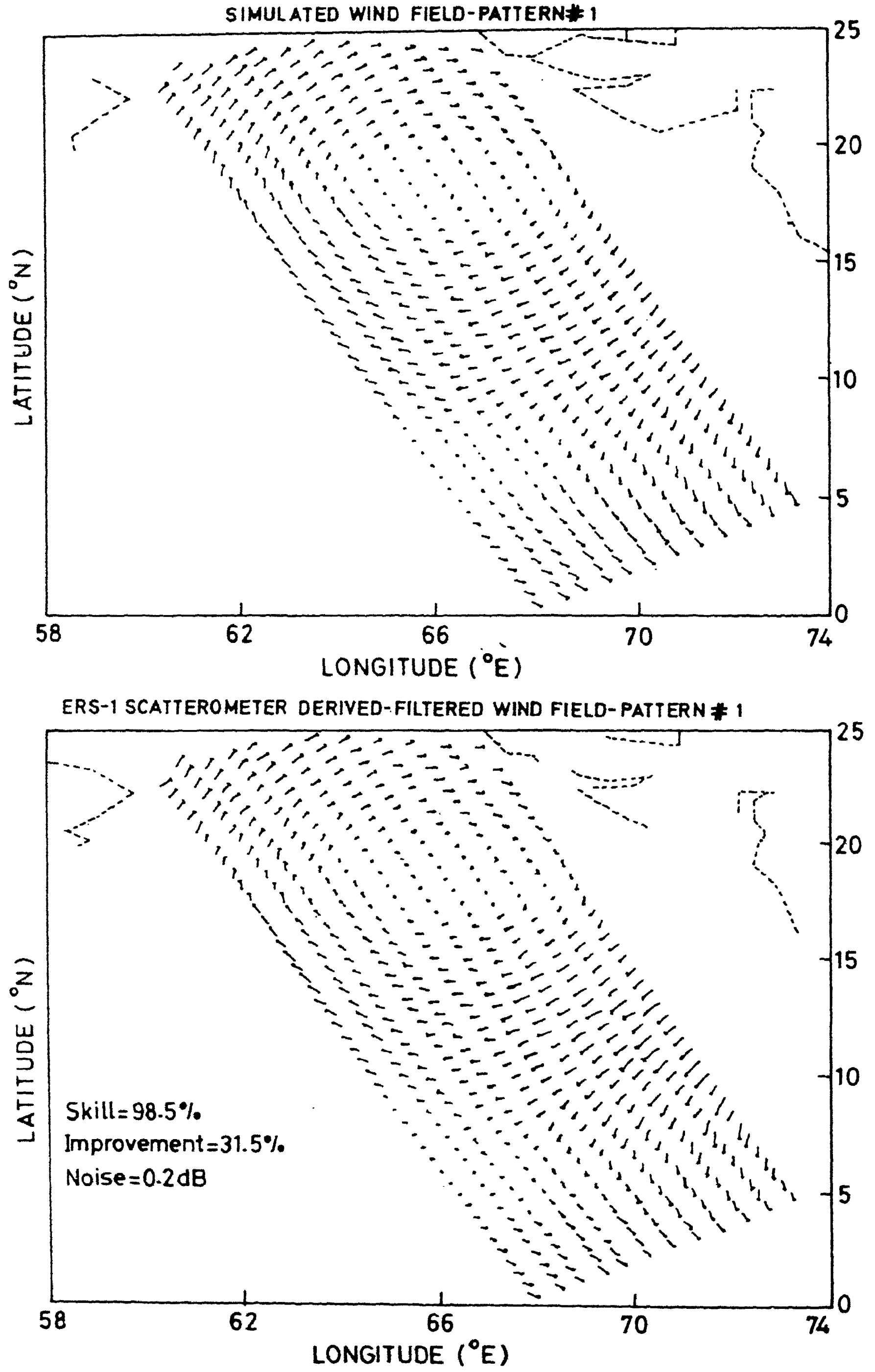


Fig 8 ERS-1 simulated results for wind vector retrieval (after Gohil 1992)[25]

radar altimeter data. Pandey[22] compared the wind speed obtained from SEASAT Scatterometer, radar altimeter and microwave radiometers. They found very interesting discrepancy and similarity which were not independent of location. Through such intercomparisons, basic questions about the physical processes responsible for the signal at the satellite and the calibration of the instruments were brought out. Woiceshyn[27] later on also noted similar discrepancy in extreme wind conditions between SASS winds and *in situ* or from pressure fields.

In addition to microwaves, which are very unique in terms of providing wind speed and wind vector, Bhandari *et al.*[28] used INSAT-VHRR images successfully to obtain wind speed from sunglint pattern of the image. Although wind speed derived from sunglint studies are limited in scope in comparison with Scatterometer and radiometers due to limited spatial coverage, nevertheless they provide an alternative "*in situ*" measurement of wind speed to compare with microwave measurements such as from sensors on board ERS-1.

Precipitation Measurements

Recently Gairola and Krishnamurthy[14] gave an algorithm for evaluating rain fall rate over oceans from a mix of satellite and surface based observations. The component of the data sets used were the outgoing Long Wave Radiation, microwave radiometric data from the special sensor microwave imager (SSM/I) and the surface rain gauge data. The method proposed by Gairola and Krishnamurty[14] assimilates these three data sets to provide improved field of global tropical rainfall. The methods are being used for initialisation of the global numerical weather prediction models and Krishnamurty's Group has noted the large impact. Similar effort is underway with Space Applications Centre and NCMRWF, Delhi.

Prior to the application of hybrid approach mentioned above, Bhandari *et al.*[29] studied precipitation measurements from INSAT images and its comparison with coastal radar at Madras in collaboration with IMD

Sea Surface Temperature

The sea surface temperature from satellite platforms are obtained from both using thermal infrared channels and microwave channels. The 3.7, 11 and 12μm channels in the infrared have been used on board TIROS-N series of operational polar orbiting weather satellites to continuously monitor global sea surface temperature for cloud free regions. The sensor is called AVHRR (Advanced Very High Resolution Radiometer) and is a five channel imaging device whose visible channels are used for cloud filtering. The two visible channels have also been used for global vegetation monitoring in terms of Normalised Difference Vegetation Index (NDVI). The next generation of AVHRR planned to be introduced in the near future will have a channel at 1.6μm which will be switchable with 3.7μm channel. This new channel will be useful for day time operation and will provide better measurements of marine aerosols. An improved configuration for SST measurements is provided by ATSR/M (Along Track Scanning Radiometer-Microwave) from ERS-1 satellite. It is expected that ATSR, which provides multiangle and multichannel measurements will provide a two fold improvement in accuracy. These sensors are providing SST to an accuracy of .5 to 1K under different conditions[30,31]. In the tropics, the problem of water vapour reduces the accuracy of SST measurements. Moreover, the single channel observations such as INSAT-VHRR are of limited use for quantitative SST determination[32].

The history of SST measurements using microwave sensors from space platform are new in comparison with IR sensors which has a long history. The SEASAT and NIMBUS-7 satellites carried a ten channel microwave radiometer whose 6.6 GHz channels have been used for sst determination, alongwith other channels for atmospheric correction[33,34]. The accuracy with these experimental satellites are 1 to 1.5 K, but under cloudy conditions also where IR measurements are not available. Much of these errors could result from the method of comparison, skin temperature versus bulk temperature, spatial average versus point measurements, lack of true coincidence etc. Moreover, the footprint size of microwave channels are quite large, (~ 150km for seasat) in comparison with IR (2-1km). Thus microwave measurements are suitable for large scale applications and are complimentary to IR measurements. Fig. 9 gives global SST obtained from SEASAT SMMR data.

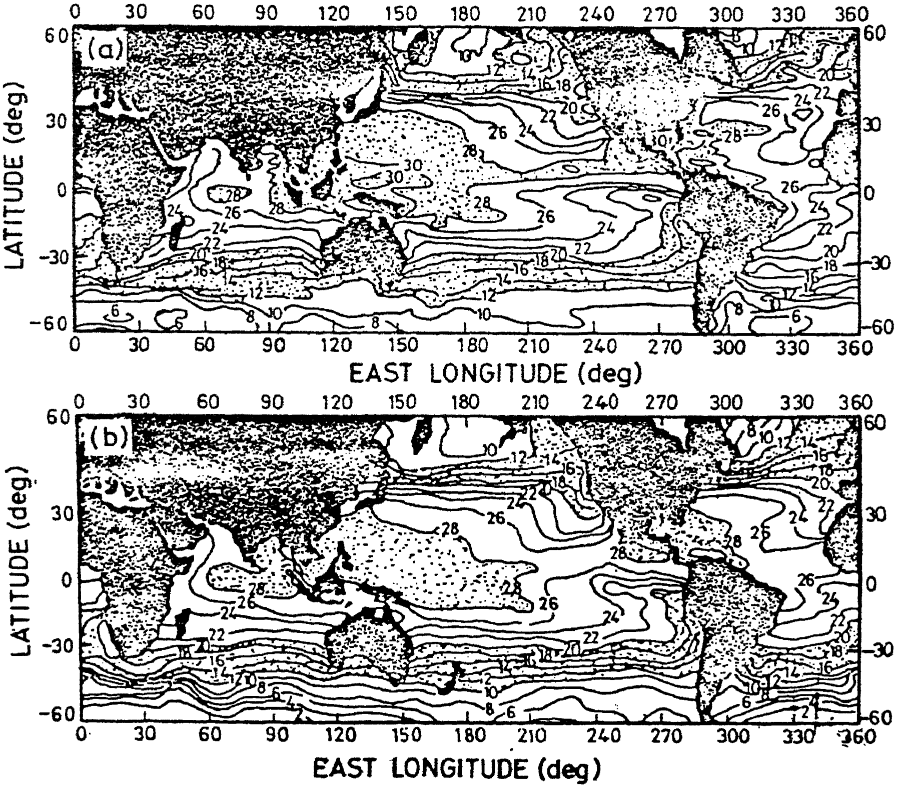


Fig 9 Global SST obtained from (a) SEASAT SMMR data for the month of August 1978 and (b) Climatological SST map for August 1978 (after Pandey and Njoku, 1982, *JPL, Private Commun*)

It is often departure from the mean SST distribution that is important rather than the distribution itself, and one of the most significant short term climatic fluctuations is also with anomalous distribution of SST in the tropical Pacific Ocean. This is El-Nino Southern Oscillation which occurs every few years and persists for about a year.

Other Studies with Satellite Data

Other interesting studies have been carried out to optimise sensors and improve the data analysis techniques[2]. In one of the applications of altimeters da-

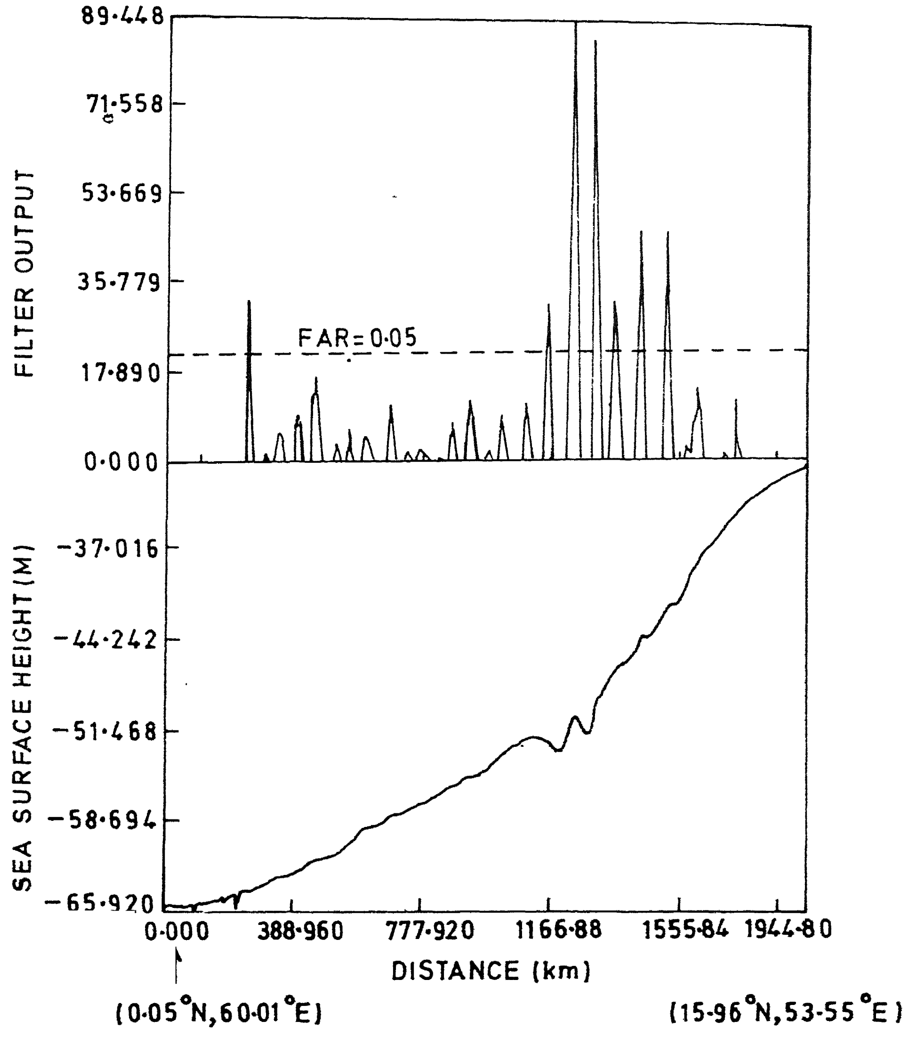


Fig 10 Detection of error seamount in the collinear track from GEOSAT Altimeter data (after Basu *et al.* 1990)[35]

ta, Basu *et al.*[35] developed the matched filtering technique to detect seamounts over the Arabian sea, Bay of Bengal, and the Indian oceans using GEOSAT altimeter data. The filter was designed using typical models for the geoid undulation signature of a sea mount and power spectral density of the background noise. The efficiency of the matched filtering technique was tested by introducing a virtual sea mount signature in the background noise. The output of the filter faithfully located the virtual seamount signature. The technique was also validated by precisely locating the charted seamounts like Somali Ridge and Error seamounts in the Arabian sea. Several new sea mounts were also predicted. These studies have found applications in Geology and fisheries, in addition to other oceanographic applications. An example of sea mount detection is given from Basu *et al.*[35] Basu and Pandey[36], did numerical experiments with modeled return echo of a satellite altimeter from a rough ocean surface and a simple iterative algorithm for the estimation of significant wave height was developed.

International Efforts

Several international programs for the study of oceans, atmosphere and hydrology have been launched in recent years, some of them are ongoing and others are at different stages of planning. The objective of listing some of these programs is to encourage scientists to direct their research keeping in mind the ongoing international large scale programs.

1. World Climate Research Programme (WCRP)
2. Tropical Oceans and Global Programme (TOGA)
3. World Ocean Circulation Experiments (WOCE)
4. International Hydrology Programme (IHP)
5. Global Energy and Water Cycle Experiments (GEWEX)
6. Joint Global Ocean Flux Study (JGOFS)
7. International Satellite Cloud Climatology Projects (ISCCP)
8. International Global Atmospheric Chemistry Programme (IGAC)
9. International Geosphere-Biosphere Programme (IGBP)

The literature on these projects is widely circulated to the scientific community. The major inputs to these large scale experiments are likely to come from satellites. In India, several institutes have taken up some of these activities, but not with enthusiam that it deserves.

Future Outlook

Looking at the major efforts undertaken by various countries notably among them, NASA, ESA, Japan, Germany, and to a modest scale by China & India, the future of application technology from space appears to be encouraging in monitoring our earth's environments. ISRO's new initiative to define an ocean satellite system in response to the Department of Ocean Development requirements is a positive step in meeting the data requirements of the specific scientific community. New instruments need to be developed taking advantage of fast progressing technology of microwave and laser systems. Optical, thermal and microwave sensors are all needed to observe different variables of the earth surface which will eventually converge in synergisms in future operational systems. The mass of data available is of itself a major focal point of the data management problem for remotely sensed data from satellites. Finally, the challenge lies in 'assimilating' these data from various sources into numerical models for prediction.

Acknowledgements

The author would like to thank Dr P K Das, FNA, Guest Editor, Atmospheric Physics, Indian National Science Academy, New Delhi and former Director General of IMD for inviting me to write this article. Thanks are also due to my various colleagues specially Mr Raj Kumar, Scientist MOD Division for helping me at various stages in the preparation of the materials presented in this paper.

References

1 P K Das *The Monsoons* National Book Trust India (1991)

1 (a) S Chandrasekhar *Radiative Transfer* Dover Publications New York USA 1960

2 P C Pandey, R K Kakar *IEEE Trans Antenna Propag* **AP-31** (1982) 136-142

3 F T Ulaby, R K Moore and A K Fung *Microwave Remote Sensing* Vol. III: *From Theory to Applications* Artech House Canton Street Norwood MA 02062 USA (1986)

4 P C Pandey, B S Gohil and T A Hariharan *IEEE Trans Geosci Remote Sensing* **GE-22** (1984) 647-655

5 R M Gairola, B S Gohil and P C Pandey *Remote Sensing Environ* **18**(2) (1985) 125-135

6 B S Gohil, A K Sharma and P C Pandey *Proc Indian Acad Sci* (*Earth Planet Sci*) **90** (1981) 263-271

7 J W Waters, K F Kunzi, R L Pettyjohn, R K L Poon and D H Staelin *J atmos Sci* **32** (1975) 1953-1969

8 C D Rodgers *Rev Geophys Space Phys* **14** (1976) 609-6245

9 P N Pathak and N Gautam *Mausam* **43** (1992) 385-394

10 P C Pandey, A K Sharma and B S Gohil *Mausam* **32** (198[illegible]) [illegible]7-22

11 P C Pandey and R K Kakar *IEEE Trans Geosci Remote Sensing* **GE-21** (1983) 208-214

12 P C Pandey *J Met atmos.Phys* **47** (1992) 165-175

13 R M Gairola, N Gautam and P C Pandey *Indian J Radio Space Phys* **21** (1992) 143-148

14 R M Gairola and T N Krishnamurty *J met atmos Phys* **50** (1992) 165-174

15 N Gautam, S Basu, R M Gairola and P C Pandey *Remote Sensing Environ* (1993)

16 Nimmi Nair, P C Pandey, B Simon and N Gautam *Int J Remote Sensing* (1993) submitted

17 W T Liu *Mon Wea Rev* **114** (1986) 1591-1602

18 W T Liu *J Geophy Res* **93** (1988) 6749-6760

19 B Simon and P C Joshi *Proc IGU Symp Hyderabad* India (1991)

20 C M Kishtwal, N Gautam, S Jaggi and P C Pandey *Tellus* (1993) *submitted*

21 C M Kishtwal, P K Pal and M S Narayanan *Proc Indian Acad Sci* (*Earth Planet Sci*) **100** (1991) 341-359

22 P C Pandey *IEEE Trans Geosci Remote Sensing* **GE-25** (1987) 670-676

23 R A Brown, W L Jones, D H Boggs, E M Bracalente, T H Guymer, D Chelton and L C Schroeder *J Geophys Res* **87**(2) 3355-3364

24 B S Gohil and P C Pandey *J Geophys Res* **90** (1985) 7307-7311

25 B S Gohil *Int J Remote Sensing* **13** (1992) 3311-3327

26 A Sarkar and Raj Kumar *Int J Remote Sensing* **7** (1986) 1369-1375

27 P M Woiceshyn, M G Wurtule, D H Boggs, L S McGoldrick and S Peteherych *J Geophys Res* **91** (1986) 2273-2288

28 S M Bhandari, W J Prakash and A K Varma Sun glint, its predictions and operations in INSAT VHRR imageries *Scient Rep Space Applications Centre* Ahmedabad (1992)

29 S M Bhandari, B M Rao, M S Narayanan, S Raghavan, D W Martin and B Auvine *Scient Rep* **ISRO-SAC-SR-27-87** Indian Space Research Organisation—Bangalore India (1987)

30 V K Agarwal and Asha Jayanti *Remote Sensing Environ* **12** (1982) 371-378

31 P J Minnet In; *Microwave Remote Sensing for Oceanographic and Marine Weather Forecast Models* (*Ed* R A Vaughan) Kluwer Academic Publishers Dordrecht London (1960)

32 R R Kelkar, V S Tewari and Sant Prasad *Mausam* **40** (1989) 13-18

33 E G Njoku and L Swanson *Mon Wea Rev* **111** (1983) 1977-1987

34 P C Pandey and Stacy Kniffen *Int J Remote Sensing* **12** (1991) 2493-2511

35 S Basu, R M Gairola and P C Pandey *Marine Geology* 14 (1990) 121-135

36 S Basu and P C Pandey *Proc Indian Acad Sci* (Earth-Planet-Sci) **100** (1991) 155-163

REMOTE SENSING AND VEGETATION

B L DEEKSHATULU AND R K GUPTA
National Remote Sensing Agency, Balanagar, Hyderabad 500 037, *India*

(*Received and Accepted* 13 *September* 1993)

The paper discusses the sensor characteristics of the existing and the forthcoming earth resources and environmental satellites useful for management and monitoring of resources, respectively. After providing a discussion on a generic basis for the remote sensing of vegetation in visible through middle infrared (IR) spectral region and on its use in agriculture and forest resources studies over Indian region, the paper emphasizes on the evolution of Vegetation Indices (VIs) concept dealing with ratio as well as orthogonal VIs. The relative sensitivity of Ratio Vegetation Index (RVI) and Normalized Difference Vegetation Index (NDVI) as a function of different phenological stages of crop growth cycle has been discussed analytically with corroborative observations based evidence. The 1.1 to >15km spatial resolution based NDVI is assuming the role of path finder data sets for Geosphere-Biosphere studies. The differential behaviour of soil component in red and near infrared reflectances as a function of crop canopy growth has been discussed to explain the Soil Adjusted Vegetation Index (SAVI). The effect of soil moisture on soil line concept especially in middle IR spectral region, concept of Weighted Difference VI(WDVI), relative advantages of middle IR and near IR reflectances in assessing water stress in vegetation, use of Integrated NDVI (INDVI) for Net Primary Productivity (NPP) and crop yield estimation, relationship of NDVI with Leaf Area Index (LAI) and Absorbed Photosynthetically Active Radiation (APAR), and limitations of VIs in accounting the vegetative processes have been discussed. The paper presents a discussion on the procedures for estimating evapotranspiration using radiation energy balance methods and the role of surface/canopy temperature obtained from satellite data in the computation of evapotranspiration. Study of processes and interlinkages among various aspects of earth systems science and incorporation of parameterization of land surface processes in Global Circulation Model (GCMs) would involve diagnostic build up of spectral information of land surface and heterogenity within a given cell of $100 \times 100\text{km}^2$ and $10 \times 10\text{km}^2$ spatial dimensions using $30 \times 30\text{m}^2$, $250 \times 250\text{m}^2$, $500 \times 500\text{m}^2$ and $1000 \times 1000\text{m}^2$ spatial resolution satellite measurements. Concepts and attempts in achieving the spatial integration to cater to the need of GCMs have been discussed. The study of processes of the earth system from space based platforms needs new initiatives and the paper also discusses about the Earth Observations System (EOS) becoming available in 1998-2015A.D. time-frame and briefly presents the capabilities of those sensors which would enhance measurement capabilities over vegetation in an exponential mode.

Key Words: Remote Sensing; Vegetation; Visible & IR Spectral Regions; Vegetation Indices; Net Primary Productivity; Global Circulation; Earth Observation System

Introduction

Remote Sensing has evolved into an important supplement to ground observations in the studies of terrestrial vegetation. Remote Sensing is the science and art of acquiring information about objects from measurements made at a distance without any physical contact with the objects. It is widely recognized that satellite remote sensing can provide an inexpensive, rapid and effective method of data collection and production of various kinds of thematic maps. The main factors which make satellite gathered data valuable in natural resources' surveys and monitoring are the synoptic view of earth resources satellites, frequent repetitive coverage which enables the observation of subtle seasonal changes, the availability of data collected in selected portions of the electromagnetic spectrum and the possibility of analysing the data both digitally and by analog.

Remote sensing of vegetation is largely connected with the measurement of its response to electromagnetic radiation, which is a continuum of radiation of different wavelengths or frequencies. The commonly used wavelength range for optical remote sensing measurements is from 0.30 to 15μm. This range is divided into several discrete/finite bands, based on the interactive response of electromagnetic radiation with various objects, and the requirement of minimum spatial dimension (picture-element: pixel) to integrate the energy in a given band to achieve the designed signal-to-noise ratio. Here, one trades-off between spectral and spatial resolution. By switching to fixed linear array self-scanning technique instead of sweep scanning technique, the integration time for the incoming spectral reflectance or signature could be increased. This resulted in higher value for signal and thus could be trade-off by decreasing the size of the pixel for a given spectral band-width and signal-to-noise ratio.

The earth resources' satellites are being primarily launched for studying and monitoring global vegetation including agriculture. In fact, the origin of LANDSAT lies in agriculture[1]. The high spatial resolution earth resources satellite data had been primarily used by the independent analysis of each spectral band image or by the flase colour compositing technique of overlaying three spectral bands to generate a composite image (discussed later). With the increased use of digital analysis, the image processing techniques had been in use to enhance the display of images by transforming the radiometric value of pixel in linear or nonlinear mode using analytical and statistical methods. Further, to club together the pixels of similar spectral signatures (in multispectral data sets) within the defined co-variance limits, statistical and mathematical morphology techniques have been used to obtain digital-analysis based classification (for different themes) in satellite data.

The concept of vegetation indices got evolved for using high spatial resolution data sets of earth resources satellites. Later on, the concept got extended for studying large scale vegetation dynamics through the use of coarse (1.1km) spatial resolution Advanced Very High Resolution Radiometer (AVHRR) data available from U.S. National Oceanic and Atmospheric Administration (NOAA) near-polar orbiting meteorological satellites. Table I gives the sensor and utility characteristics of 4-bands Multi-Spectral Scanner (MSS) and 7-bands Thematic Mapper (TM) on-board LANDSAT (USA) satellite, Linear Imaging Self Scanning Sensors (LISS-I & II) on-board Indian Remote Sensing Satellie (IRS), and Panchromatic Linear Array (PLA) and Multispectral Linear Array (MLA) sensors on-board SPOT (France) satellite. Table II gives the sensor and utility characteristics of forthcoming IRS-1C Landsat-6 launched but

Table I

Characteristics and utility potential of existing remote sensing satellites

Year of Launch	Landsat 1, 2, 3 1972, 1975, 1978	IRS 1A, 18 1988, 1991		Landsat-4,5 (TM & MSS) 1982, 1984	Primary Utility	SPOT-1,2 1986, 1990
Spectral Bands: (Wavelength in Micro meters)					Cartographic Works	0.51-0.73 (Panchromatic: PLA)
	MSS	0.45-0.52 (Band 1)		0.45-0.52 (TM1)	Coastal water mapping; Soil/ Vegetation & Deciduous/ Coniferous differentiation	
						Multispectral (MLA)
	0.5-0.6 (Band 4)	0.52-0.59 (Band 2)		0.52-0.60 (TM2)	Green reflectance from healthy vegetation	0.50-0.59 (Band 1)
	0.6-0.7 (Band 5)	0.62-0.68 (Band 3)		0.63-0.69 (TM3)	Chlorophyll absorption of plant species differentiation	0.61-0.68 (Band 2)
	0.7-0.8 0.8-1.1 (Bands 6&7)	0.77-0.86 (Band 4)		0.76-0.90 (TM4)	Biomass surveys and water bodies delineation	0.79-0.89 (Band 3)
				1.55-1.75 (TM5)	Snow/Cloud differentiation; Vegetation moisture measurement	(two identical cameras of 60 km swath width with 3 km overlap)
				10.40-12.50 (TM6)	Plant heat stress measurement and general thermal mapping	
				2.08-2.35 (TM7)	Hydrothermal mapping	
Spatial Resolution	80m	LISS-I LISS-II	72.5m 36.25m	30m TM6-120m		MLA 20m PLA 10m
Swath Width	185km	LISS-I LISS-II	148km 74km	185Km		117km
Radiometric Resolution	(7 bits for MSS-7)	7 bits		8 bits		8 bits
Equatorial Crossing Time	09:30 AM	10:08 AM		09:30 AM		10:30 AM
Repeat Cycle	18 days	22 days		16 days		26 days

Table II

Details of forthcoming remote sensing satellites

Name of Satellite:	IRS-1C (India)			SPOT-4 (France)	
Launch Year	1994			1996	
Spectral Bands (in micro meters)	PAN	LISS-III	WIFS	PLA	MLA
	0.50-0.70	0.52-0.59		0.51-0.73	0.50-0.59
		0.62-0.68	0.62-0.68		0.61-0.68
		0.77-0.86	0.77-0.86		0.79-0.89
		1.55-1.70			1.52-1.75
Spatial Resolution	< 10m	23.4m	188m	10m	20m
Altitude (Km)		817			832
Radiometric Resolution	6 bits	7 bits	7 bits		8bits
Swath (Km) (Nadir)	70.5	141	780		117
Repeat Cycle (days)	5	24	5		26

WIFS : Wide Image Field Sensor; IRS : Indian Remote Sensing Satellite; PAN: Panchromatic;
LISS : Linear Imaging Self Scanning; PLA: Panchromatic Linear Array
MLA : Multispectral Linear Array

Table III

Characteristics of advanced very high resolution radiometer (AVHRR) on-board NOAA polar orbiting meteorological satellites and its utility

Band Number	Spectral Region (microns)	Primary Utility
1	0.55-0.9 (Tiros-N) 0.58-0.68 (NOAA-6 onwards)	Day time cloud, snow, ice and surface Mapping
2	0.725-1.1	Surface Water Delineation, Vegetation/ Agriculture Assessment, Range Surveys and Location of Water bodies
	3.55-3.93	Night time Sea Surface Temperature and Cloud Mapping; Land mark extraction, Forest Fire Monitoring and Volcanic Activities
4	10.5-11.5 (till NOAA-6) 10.3-11.3 (NOAA-7 onwards)	SST, day/night cloud mapping, Soil Moisture/Evapotranspiration, Volcanic Activities Monitoring
5	11.5-12.5 (NOAA-7 onwards)	SST, day/night cloud mapping, Surface Temperature

not traceable and SPOT-4 satellites. Table III gives the sensor characteristics and utility of NOAA/AVHRR. The sensor systems aboard different satellite platforms could be used complementarily with one another to extract more information e.g., studies have shown such complementary nature behaviour between the NOAA/AVHRR and the Landsat[2], the Landsat and the IRS[3], the NOAA and the IRS etc.

It is now believed that earth has to be studied in totality through systems approach and International Geosphere-Biosphere Programme (IGBP) has provided an initiative in this direction. Mission To Planet Earth (MTPE) has been initiated by US NASA for understanding various processes of earth systems science through Earth Observations System (EOS) during 1998-2015A.D. The NOAA meteorological and earth resources satellites shall continue to contribute to monitoring and resource management requirements of the planet earth, respectively. With this in view, the paper discusses, in detail the concept and evolution of Vegetation Indices (VIs) which are assuming the role of least common denominator for resource management, monitoring and process-understanding studies. With the initiation of integrated geosphere-biosphere studies under IGBP, the use of VIs and computation of evapotranspiration have assumed emerging roles at local through regional to global spatial scales. Similarly, nesting of temporal scales is needed to study the role of vegetation in geosphere-biosphere studies. Finally, the paper gives a brief account of EOS and of those on-board instruments which have relevance for vegetation studies.

Vegetation Response

Fig. 1 gives typical spectral response characteristics of green vegetation[4]. In the visible region chlorophyll absorption dips are approximately centered at 0.45 and 0.65μm. Absorption due to chlorophyll masks the absorption characteristics of yellow pigments in 0.45μm band and these (carotenes and xanthophylls) become dominant during senescence and are the primary cause for the yellow colouration of tree foliage in the fall season. The pigment effects are feeble in near and middle infrared (IR) regions. In near IR (NIR), the Bi-Directional Reflectance (BDR) is related to the degree of scattering that occurs as a direct consequence of discontinuities in refractive index within the leaf[5]. The internal structure of leaf is different for different vegetation and this difference gets noticed in NIR. The reflectance and transmission in NIR are typically of 45 to 50 per cent magnitude leaving absorption to be at around 5 per cent level. This high transmission permits the BDR to contain additive reflectance from the multiple layers of leaves in the canopy. The weak water absorption bands near 0.96 and 1.1μm also contribute significantly to absorption in multiple layers of leaves. Physical and biological basis for vegetation reflectances in visible and NIR spectral regions has been discussed by Knipling[6]. Water has strong absorption bands at 1.45, 1.95 and 2.5μm in middle IR[7]. The 1.55-1.75 and 2.08-2.35μm bands of Landsat TM are sensitive to crop moisture content or leaf air space volume[8].

The images in each of these bands could be used individually and also by compositing three bands by assigning one of the three primary colours to each

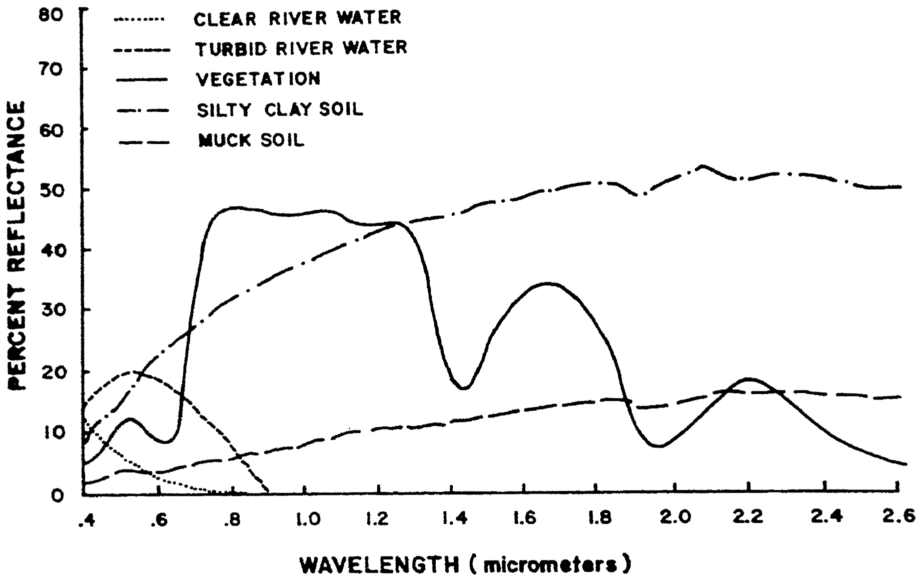


Fig 1 Spectral reflectance of basic land cover types

band. In standard False Colour Composite (FCC) Red, Green and Blue colours are assigned to MSS bands 7, 5 and 4 or TM/IRS-LISS bands 4, 3, 2 or SPOT/MLA bands 3, 2, 1 respectively. The NIR band has maximum reflectance from vegetation (Fig. 1) and assigning of red filter to it while making FCC results in the dominance of red colour for vegetation, deeper the tone higher the green biomass, in general. Similarly, appearance of shallow and deep water in light and dark blue colour (as blue is assigned to shorter wavelengths where water has maximum BDR) and the colour tones of various other objects in the FCC could be explained using spectral response curves such as in Fig. 1.

These FCCs and/or digital analysis techniques had been used for studying forest resources[9-18], grasslands[19], estimating biomass[20-21], forest fire monitoring[22], agriculture[23] and for estimating crop acreage[24-29] over Indian region. Medhavy *et al.*[30] have studied the accuracy of acreage estimation and Potdar *et al.*[31] have studied the relative performance of Landsat TM and IRS-LISS-I for acreage estimation. A review on the use of space technology in forest management is available in Rao[32]. Remote Sensing is being operationally used for mapping and monitoring of forest and agriculture. Inclusion of mid IR(MIR) bands of TM with visible and NIR bands brought about better discrimination between rainfed rice crop and hill region vegetation[33]. Pereira and Setzer[34] have shown the potential of 3.55-3.93μm AVHRR MIR band in detecting forest fire while Nelson and Horning[35] have used AVHRR data for estimating forest area in Madagascar.

Vegetation Indices

The vegetation or biomass indices basically utilize the differential aspects of vegetation reflectance in Red (or visible in case of AVHRR) and NIR. As MSS was the first sensor, the basic discussion on VIs will be referring to it. As the plant grows healthy the NIR reflectance would increase and the red (or visible) reflectance would decrease (Fig. 1). This opposite nature of variation could be algebraically enhanced and spectrally compressed by taking ratio and difference of reflectances in these two bands as given by

Ratio Vegetation Index (RVI) = Near IR(MSS7)/Red(MSS5) ... (1)

Difference Vegetation Index (DVI) = Near IR(NIR) − Red(R) ... (2)

To take care of seasonal variations in illumination conditions and atmospheric effects, the DVI could be normalized by the sum of NIR and Red (R) to give

Normalized Difference Vegetation Index (NDVI) = (NIR − R)/(NIR + R) ... (3)

In case of MSS, the band 6 (MSS6) could also be used in place of band 7 (MSS7) and MSS 6 refers to rising portion of plant response in NIR while MSS 7 refers to near flat portion of NIR (Fig. 1). To take care of negative values encountered in NDVI over cloud, snow and water zones of satellite images, a value of 0.5 was added by Rouse *et al.*[36], and the square root of this, to reduce variance, was termed as Transformed Vegetation Index (TVI) given by

$\mathrm{TVI} = (\mathrm{NDVI} + 0.5)^{1/2}$ ... (4)

Perry and Lautenschlager[37] developed another form of TVI given by

$\mathrm{TVI} = [(\mathrm{NDVI} + 0.5)/\{\mathrm{ABS}(\mathrm{NDVI} + 0.5)\}] * [\mathrm{ABS}(\mathrm{NDVI} + 0.5)]^{1/2}$ ... (5)

Here, ABS denotes the absolute values and occurrence of 0/0 is set to 1.
To take care of lower intensity levels in NIR illumination, following new vegetation indices were developed.

DVI[38] = 2.4 (MSS7) − MSS5 (6)

Ashburn Vegetation Index (AVI)[39] = 2.0 (MSS 7) − MSS 5 . (7)

Difference − Difference Vegetation Index (DDVI)[40] = (2 * MSS7 − MSS6) − (MSS5 − MSS4) . (8)

The bispectral plot between R and NIR for an uncultivated field would give a near straight line, called soil line. As the crop grows over it the crop response curve would move in a direction perpendicular to this soil line and would assume tasseled cap shape during the peak vegetative phase. Using this behaviour, the concept of Perpendicular Vegetation Index (PVI7 and PVI6 refer to the use of MSS 7 or MSS 6 as NIR band, respectively) got developed[38]. PVI7 and PVI6 are given by

$\mathrm{PVI7} = [(0.355\,\mathrm{MSS7} - 0.149\,\mathrm{MSS5})^2 + (0.355\,\mathrm{MSS5} - 0.852\,\mathrm{MSS7})^2]^{1/2}$ ... (9)

$\mathrm{PVI6} = [(0.498\,\mathrm{MSS6} - 0.457\,\mathrm{MSS5} - 0.498)^2 + (0.498\,\mathrm{MSS5} - 0.543\,\mathrm{MSS6} + 2.734)^2]^{1/2}$ ... (10)

Shibayama *et al.*[41], using ground based measurements, have shown that though the PVI is subject to bidirectional effects, it contained more useful information about wheat canopies at Leaf Area Index (LAI) > 2 than does the NDVI.

The orthogonal vegetation index was developed by Kauth and Thomas[42] by applying a heuristic linear transformation in the 4-dimensional data space provided by the Landsat MSS measurements over vegetation for different soil types and conditions. These Tasseled Cap Transformations (TCT) were derived through rotation of principal components computed using covariance matrices which were arrived at using images over bare soil and different crop growth stages. The first through fourth components of TCTs were named as Greenness Vegetation Index (GVI), Brightness Soil Index (BSI), Yellowness Stuff Index (YSI) and Non-Such Index (NSI) given by

$$\mathrm{GVI} = -0.28317\,\mathrm{MSS4} - 0.66006\,\mathrm{MSS5} + 0.57735\,\mathrm{MSS6} + 0.38833\,\mathrm{MSS7} \quad \ldots(11)$$

$$\mathrm{BSI} = 0.33231\,\mathrm{MSS4} + 0.60316\,\mathrm{MSS5} + 0.67581\,\mathrm{MSS6} + 0.26278\,\mathrm{MSS7} \quad (12)$$

$$\mathrm{YSI} = -0.89952\,\mathrm{MSS4} + 0.42830\,\mathrm{MSS5} + 0.07592\,\mathrm{MSS6} - 0.04080\,\mathrm{MSS7} \quad \ldots(13)$$

$$\mathrm{NSI} = -0.01594\,\mathrm{MSS4} + 0.13068\,\mathrm{MSS5} - 0.45187\,\mathrm{MSS6} + 0.88232\,\mathrm{MSS7} \quad (14)$$

Hay *et al.*[43] further combined the first two TCT components to compute Greenness Above Bare Soil (GABS) given by:

$$\mathrm{GABS} = \mathrm{GVI} - 0.09178\,\mathrm{SBI} + 5.58959 \quad \ldots(15)$$

The TCT concept was extended to six reflective spectral channels (0.45 − 0.52, 0.52 − 0.60, 0.63 − 0.69, 0.76 − 0.90, 1.55 − 1.75 and 2.08 − 2.35μm) of Landsat TM by Crist and Cicone[44] and on comparison with MSS, it was inferred that at least one or two additional components of information beyond the first two components (greenness and brightness) of MSS data were feasible and these were due to new information on plant water content available from mid infrared bands (1.55-1.75 and 2.08-2.35μm) of Landsat. The first two TCTs arrived at using Landsat TM bands 2,3 and 4 were found to be similar with respect to first two TCTs obtained using 4 bands MSS data. Thus, it could be inferred that the 0.9-1.1μm region leads to no discernible loss of information in TM data sets as compared to MSS data sets.

Cicone and Metzler[45] applied TCT concept to visible (AVHRR1) and NIR (AVHRR2) channels of NOAA/AVHRR and obtained

$$\text{AVHRR Greenness Index (AGI)} = 0.89\,\text{AVHRR1 Radiance} - 0.45\,\text{AVHRR2 Radiance} \quad \ldots(16)$$

$$\text{AVHRR Brightness Index (ABI)} = 0.45\,\text{AVHRR1 Radiance} + 0.89\,\text{AVHRR2 Radiance} \quad \ldots(17)$$

Perpendicular Vegetation Index (PVI) developed by Richardson and Wiegand[38] requires the reflectance of bare soil which often is not known in order to offer a more precise correction for soil background. Fig. 2 depicts[46] isolines (lines of constant vegetation amounts) for ratio (RVI and NDVI) and orthogonal (PVI) indices. The isolines of ratio vegetation indices appear as lines of increasing slope diverging out from the origin while PVI lines appear parallel to soil line. In red, most of the incoming irradiance is absorbed by the vegetation canopy and thus, soil signatures are limited by irradiance received directly from Sun and sky through gaps in canopy. In case of NIR, due to nearly 50% transmission through plant leaves, the soil is additionally irradiated by this energy as well as what is reflected by the ground side of upper leaf from the impinging upwardly reflected (≈ 45%) energy from the upper layer of lower leaf. Under low canopy condition there is not enough vegetation to impart this additional irradiance to soil in NIR while under high vegetative conditions not enough signal from the soil will emerge out of canopy. Thus, the differential behaviour of NIR and red irradiances of soil becomes important at about 50% vegetative cover condition. This makes the convergence of isolines to a common point a function of LAI.

Point 'A' in Fig. 2 refers to partial canopy over dry soil background. Once the soil becomes wet, the bispectral reflectance would shift to location 'B' and 'C' on the RVI/NDVI and PVI isolines, respectively, to maintain the same VI value under wet as well as dry soil conditions. By conducting experiments over a series of cotton and range grass canopies under different types and conditions of soil, Heute[46] found vegetation isolines to possess NIR-R wavelengths slope in between those of ratio and orthogonal-based (PVI) isolines as indicated by point 'D' in Fig. 2. As first order approximation, this could be taken as

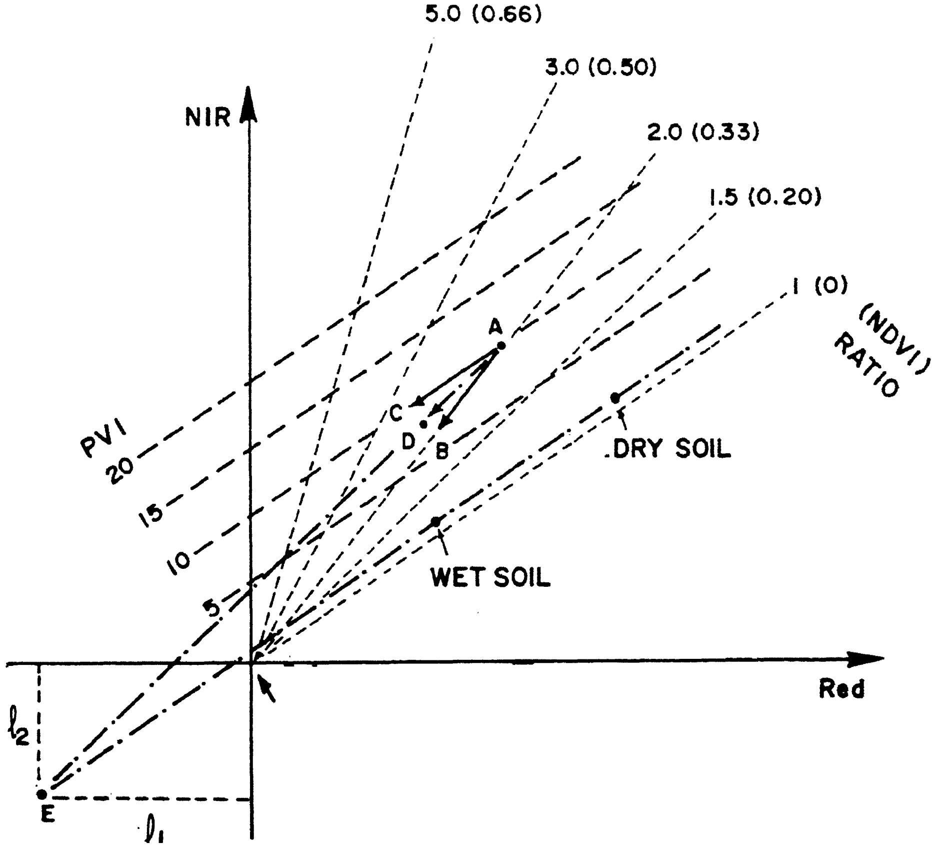


Fig 2 Vegetation spectra isolines in NIR-red wavelength space as predicted by the ratio-normal-

shifting of origin to a point where intermediate densities of vegetation converge with the soil line. The shifting of origin refers to redefining of this Soil-Adjusted Vegetation NDVI Index (SAVI) as

$$\mathrm{SAVI} = \{(\mathrm{NIR}-\mathrm{R})+(l_2-l_1)\}/\{(\mathrm{NIR}+\mathrm{R})(l_1+l_2)\} \quad \ldots (18)$$

Under the assumption that soil line has a slope close to l, l_1 could be taken as nearly equal to l_2 and this reduces eq. (18) to

$$\mathrm{SAVI} = (\mathrm{NIR}-\mathrm{R})/[(\mathrm{NIR}+\mathrm{R})+L], \quad \ldots (19)$$

where $L = l_1 + l_2 = 2l$

In order to keep SAVI bounded within the theoretical limits of -1 and $+1$ (as in case of NDVI), Heute[46] applied a multiplication factor of $(1+L)$ to eq. (19) and redefined SAVI as:

$$\mathrm{SAVI} = [(\mathrm{NIR}-\mathrm{R})/\{(\mathrm{NIR}+\mathrm{R})+L\}]^*(1+L) \quad \ldots (20)$$

and opined that a value of 0.5 for L is a good approximation, in general. Baret *et al.*[47] further improved the SAVI to Transformed Soil Adjusted Vegetation Index (TSAVI). As the rate of change of soil response in R − NIR bispectral plot is nearly constant, Li *et al.*[48] have proposed the concept of indices based on the second derivative reflectances to minimize the effect of variations over the soil surface.

Baret *et al.*[49], analyzing the concept of soil line using simple radiative transfer model splitted the BDR into its single and multiple scattering components, and have shown that slope of soil line in R-NIR bispectral plot corresponds to the ratio of single scattering albedos while intercept of the soil line has origin in the difference of multiple scattering components. In the water absorption band related to MIR region, the change in soil moisture content will introduce scatter in soil line, in linear agreement with soil reflectance.

Spectral vegetation indices in the visible and NIR wave-bands have been shown to be broadly diagnostic of geography and seasonality of vegetation. The NOAA/AVHRR NDVI has been used for vegetation monitoring[50-56], primary productivity studies[57,58], estimating dry matter production[59], and for land cover mapping/monitoring[60-63]. Tucker *et al.*[64] have related atmospheric CO_2 with vegetation index. AVHRR NDVI has been operationally used for drought monitoring over India[65]. The district level averages of AVHRR NDVI and RVI have been related with Growing Degree Days (GDDs) of mean and maximum air temperatures, over the whole of Crop Growth Cycle (CGC), for the predominant wheat sowing districts of Haryana and Punjab with 98 to 99% statistical significance level[66,67]. The land applications potential and standardised methodology for the use of NOAA/AVHRR NDVI are available in the IGBP[68].

NDVI could be expressed in terms of RVI by the equation

$$\mathrm{NDVI} = (\mathrm{RVI}-1)/(\mathrm{RVI}+1) \quad \ldots (21)$$

Differentiating NDVI with respect to RVI gives $[d(\mathrm{NDVI})/d(\mathrm{RVI})] = 2/(\mathrm{RVI}+1)^2$

$$\ldots (22)$$

This shows that NDVI and RVI are non-linearly related and are thus expected to provide different levels of sensitivity for different stages of CGC. As vegetation indices are related to LAI, these represent the integrated effect of crop growth history till the date of measurement and thus takes care for the resultant effect of nonlinear processes—an important reason at the credit of satellite data based biomass/vegetation index estimation.

In order to do relative sensitivity analysis of RVI and NDVI, it can be assumed that errors caused by variuous factors are e_1 and e_2 in the reflectances for R and NIR bands, respectively. Thus, the modified versions of equations denoted with RVI' and NDVI' for RVI and NDVI respectiely, could be rewritten as

$$\mathrm{RVI}' = (\mathrm{NIR} + e_2)/(\mathrm{R} + e_1) \qquad \ldots (23)$$

or

$$\mathrm{RVI}' = \mathrm{RVI}[(1 + e_2/\mathrm{NIR})/(1 + e_1/\mathrm{R})] \qquad \ldots (24)$$

$$\mathrm{NDVI}' = [(\mathrm{NIR} - \mathrm{R}) + (e_2 - e_1)]/[(\mathrm{NIR} + \mathrm{R}) + (e_2 + e_1)] \qquad \ldots (25)$$

or

$$\mathrm{NDVI}' = \mathrm{NDVI}[[1 + \{(e_2 - e_1)/(\mathrm{NIR} - \mathrm{R})\}]/[1 + \{(e_2 + e_1)/(\mathrm{NIR} + \mathrm{R})\}]] \qquad \ldots (26)$$

The relative sensitivity of the RVI and NDVI taken as a function of green biomass and can be categorized in two broad situations:

Situation 'A'

This situation would be experienced during initial crop growth and senescence stages.
Here, the values of R and NIR will be comparable and it can be assumed that

$$(e_2/\mathrm{NIR}) \approx (e_1/\mathrm{R}) \qquad \ldots (27)$$

The incorporation of this condition into eqs. (24) and (26) gives

$$\mathrm{RVI}' \approx \mathrm{RVI} \qquad \ldots (28)$$

$$\mathrm{NDVI}' = \mathrm{NDVI}\,[1/\{1 + \{(e_2 + e_1)/(\mathrm{NIR} + \mathrm{R})\}\}] \text{ for } (e_2 - e_1) < < (\mathrm{NIR} - \mathrm{R}) \qquad \ldots (29)$$

Thus, NDVI' would be slightly less than the actual NDVI, This emphasizes the value of using the RVI in the initial crop growth and senescence stages.

Situation 'B'

This situation would be observed during jointing, heading, flowering, milking and dough stages of the CGC. Here $\mathrm{NIR} > > \mathrm{R}$ and $\mathrm{NIR} > > e_2$. Application of these conditions would modify eqs. (24) and (26) to

$$\mathrm{RVI}' = \mathrm{RVI}\,[1 + (e_1/\mathrm{R})] \qquad \ldots (30)$$

$$\mathrm{NDVI}' = \mathrm{NDVI}\,[1/\{1 + \{(e_2 + e_1)/(\mathrm{NIR} + \mathrm{R})\}\}] \qquad \ldots (31)$$

There will be a decrease in the value of R as the crop gets into these crop growth stages and this would make the value of (e_1/R) non-zero. Thus, RVI' would certainly be more than RVI. As $(\mathrm{NIR} + \mathrm{R}) > > (e_2 + e_1)$, and

$(e_2 - e_1) << (\mathrm{NIR} - \mathrm{R})$ the NDVI′ and NDVI would be nearly comparable. Thus, the use of the NDVI as compared to the RVI would be better under this situation.

The observations based evidence for the inferences of the above analytical treatment using ground based spectral measurements over soybean canopy have been given by Gupta *et al.*[69] while extension of this to 1.1km spatial resolution data based aggregation over the major wheat growing districts is available in Gupta[70] and Gupta *et al.*[71]. While inferring about relative sensitivity of AVHRR RVI and NDVI over the district level aggregation, the preprocessing error was minimized by applying geometric correction to visible (substitutes the red band of earth resources' satellites) and NIR channels of AVHRR, towards their use for RVI and NDVI computation, and by selecting ground control points in the near proximity of study area[72]. Further only those AVHRR images were used in statistical aggregation over the district where the AVHRR scan angle limits for the district boundaries were much below 38 degree[73,74].

NDVI is commonly preferred due to partial elimination of the dependence upon sun-target-sensor geometry. If the anisotropic behaviours of reflectance in red (R) and NIR were similar, i.e., of the change of R and NIR with geometry were proportional, NDVI would be invariant with respect to the variations in viewing and illumination angles. However, the two reflectances have some differences in their bidirectional distributions and therefore, NDVI is slightly geometry dependent. In the calculation of NDVI values, the reflectance factor of the reference channel (sometimes referred to as the 'equivalent albdeo') need to be periodically updated because of changes in the sensor sensitivity with time[75]. A comparison of the NDVI calculated from reflectance or albedo would be systematically higher then the values obtained from the count values[76] and also offer more realistic year to year comparison of NDVI values.

Changes in sun angle results in variation of path radiances due to atmospheric scattering and absorption (Rayleigh, water vapour, aerosol, ozone, mixture of gases etc.), thus, having impact on the NDVI at ground level. Atmospheric effects (clouds, water vapour, aerosols) reduce the contrast between the visible and the near IR reflections, thus, usually decreasing the vegetation indices. After elimination of all causative factors responsible for low NDVI values over the southern latitudes of India, Jayaseelan and Chari[77] attributed this to the effect of aerosols after the explosion of volcano Mt. Pinatubo. They observed more reduction in the digital number value of the data of band 2 than those in band 1. The necessity for atmospheric corrections are dictated not only by the different geometry of observations but also by the global and seasonal distribution of water vapour and other atmospheric constituents. For inter-comparison of biospheric productivity at global scale, it is important to develop some simple atmospheric corrections based on measurements in relevant spectral bands; the dependency on results of radiative transfer modeling need to be reduced while implementing atmospheric corrections in operational mode. Difference of 10.3-11.3 and 11.5-12.5μm AVHRR thermal infrared channels, known to be sensitive to atmospheric water vapour amount, could be used to correct the near IR band data.

The maximum value compositing is believed to take care of most of the atmospheric effects (clouds, aerosols, water vapour etc.). But Deering and Eck[78] observed higher NDVI values in hazy conditions than on clear sky days, for full coverage under soybean canopy. To reduce the probability of cloud contamination, original weekly composites of GVI (Global Vegetation Index: NDVI with spatial resolution >15km) are sometimes further composited over a longer periods (3 or 4 weeks). Smoothening of the vegetation index time profiles by statistical filters also could be employed. This would only reduce the noise, but do not eliminate cloud contamination altogether, which decides value of the VIs. Extended periods of compositing though increase the probability of a clear observation but loose temporal resolution so that short term vegetation changes are not observed. Therefore, a straight forward way to select the clearest observation is by taking maximum of vegetation indices over a period of time. The time compositing procedures are based on this principle.

Nature of both the reflectance anisotropy and its asymmetry with respect to nadir is different for the atmosphere and for the surface. In the atmosphere, path length is of crucial importance with minimum atmospheric effect at the nadir. Anisotropy of the surface depends upon the state of vegetation, angle distribution within canopy and resultant shadowing. Since vegetation is highly anisotropic the bidirectional surface models will have to be taken into account in radiative transfer modeling. Duggan and Saunders[79], attributed the variability of AVHRR data across a scan line mainly to the anisotropy of the surface.

A biophysical explanation of the relation between vegetation indices and observable vegetation phenomena is still subject to much discussion. Progress in moving from qualitative inspections to quantitative assessment of global scale vegetation dynamics with remotely sensed observations requires a more complete biophysical understanding of the measurements. However, all the VIs suffer from the drawback that mathematical description of the relationship of such an index to crop characteristic description such as LAI is empirically derived and viewed differently from worker to worker.

For example, the crop growth models describe the relation between the physiologial process in plants and the environmental factors such as solar radiation, temperature as well as water and nutrient availability. The LAI, is a main driving variable in many crop growth models, designed for yield prediction[80]. Remote sensing has the potential of providing information about the actual status of the crop e.g., in terms of LAI, which would improve crop growth models.

Clevers[81] developed a simplified and semi-empirical reflectance model for estimating LAI of a given canopy (CLAIR). In the CLAIR model, he applied Weighted Difference VI (WDVI), which was shown to be comparable to the greenness index for the 2-dimensional case of the PVI.

$$\mathrm{WDVI} = \mathrm{R_{ir}} - C\mathrm{R_r} \qquad \ldots (32)$$

$$\text{where } C = \mathrm{R_{s,ir}}/\mathrm{R_{s,r}} \qquad \ldots (33)$$

Here $\mathrm{R_r}$ and $\mathrm{R_{ir}}$ denote total measured reflectance in red and near IR over crop respectively, $\mathrm{R_{s,r}}$ and $\mathrm{R_{s,ir}}$ denote red and near IR reflectance of bare soil, respectively. Here, a 'corrected NIR reflectance' is calculated by subtracting the contribution of the soil in line of sight from the measured reflectance of the composite canopy-soil scene. Clevers[81] related WDVI and LAI through

$$\mathrm{LAI} = (-1/\alpha)\ln(1 - \mathrm{WDVI}/\mathrm{WDVI}_\infty) \qquad \ldots (34)$$

where α is a combination of extinction and scattering coefficients, describing the rate with which the function of the above equation runs into an asymptotic value, and WDVI_∞ as the asymptotically limiting value for the WDVI. Parameters α and WDVI_∞ have to be estimated empirically from a training set, but they would have a physical basis, when fitted in the equation. The relationship between LAI and WDVI had been found to be less sensitive to soil reflectance, ratio of diffuse to total irradiance, solar zenith angle, and hot spot size[82]. It was sensitive to chlorophyll content in 10-20μg cm^{-2} range and to leaf inclination angle.

In drawing the inference of vegetation status from measurements in the visible and NIR in the presence of variable soil reflectance, Price[83] presented two cases depending upon the spatial variability within the instrumental field of view. The first one being, 'Field case', assumes a spatially uniform vegetation canopy (as in the case of data from IRS or Landsat satellites) and the second one being the 'Mixed pixel case', applies for spatial inhomogenity on a large scale (as in the case of AVHRR) observations of agricultural areas. In the first case, a value of LAI may be obtained from the reflectance measurements in the visible and NIR. In the second case, a fractional cover, *f*, corresponding to the dense vegetation may be obtained from 2 measurements and for this he introduced a leaf vegetation index V_1, having limits 0 and 1 and showed that for a thin canopy of LAI $<<1$, V_1 was to be equal to the vegetation fraction, *f*.

NDVI though sensitive at low cover values, is often a poor indictor of real-time photosynthetic fluxes and tends to saturate at high LAIs. Gamon *et al.*[84] therefore, attempted to conceptualize and develop an index, what is known as the Physiological Response Index (PRI), which reveals short term changes in actual photosynthetic efficiency, particularly in well developed canopies, where NDVI is least sensitive.

The spectral-biophysical functional relations obtained through remotely sensed data need to be validated across locations and environments. But among-site comparisons are even more difficult because of variations in sun zenith angle (latitude, sowing date and time of day for observations), soil surface conditions, instruments and measurement techniques, canopy architecture (leaf angle, canopy openness, height) and cultural practices (row spacing, plant population, fertilizers). To provide the needed relationships, the Spectral – Agronomic – Multisite – Multicrop Analysis (SAMMA) was initiated[85]. Under SAMMA, reflectance data from handheld and boom-mounted spectro-radiometers and agronomic or biophysical plant measurements have been pooled across locations for uniform analyses for the crops—wheat, corn, sorghum, soybean and *Alfalfa*[85]. Such experimens need to be conducted for the agroclimatic zones of the India to enable comparisons of cropping production efficiencies using spectral and agronomical parameters in a synergistic manner.

Using NOAA – AVHRR data it will be difficult to study a specific crop, unless the specific cropped area is of longer expanse, since the spatial resolution of AVHRR is much larger than the average field size. The obtained NDVI temporal variation mapped at 1km (and consequently at 4km) always corresponds to a mixed response of bare soil, grass land, urban areas, forest covers, different crops at different phenological stages etc. This kind of situation is more often a fact under Indian conditions, where the average size of land holdings is very small (≈ 0.2ha).

Baret *et al.*[86] have shown the complementarity of middle infrared (MIR) spectral information with that of near infrared reflectance for monitoring wheat canopies. Biological and optical properties of plant leaves largely determine the spectral reflectance from plant canopy. In the 1.3-2.5μm (MIR) spectral interval, liquid water content of leaves largely controls the resulting spectral reflectance.

Tucker[87] analysed the relationships between predicted spectral reflectance and equivalent water thickness with a model simulation and concluded that 1.55-1.75μm (TM5, one of the MIR band of Landsat TM) was the best band for monitoring canopy water status from space platforms. At the canopy level, reflectance spectra depends not only on optical properties of leaves, but also on the background soil spectra and the canopy geometry. The canopy mid-infrared response provides valuable information on crop discrimination, biomass and plant water content, although higher degree of water stress can be more easily detected in the near-infrared than in mid-infrared due to sensitivity to changes in canopy leaf geometry and canopy cover caused by wilting. Under field conditions, changes in canopy geometry induce a larger effect in the optical properties of leaves.

Under waterstress and nitrogen deficiency-free conditions Baret *et al.*[86] concluded that 1.560-1.685μm (similar to TM5) contained more information than the other mid-IR spectral bands (1.660-1.695μm and 2.030-2.235μm), and canopy reflectance in mid IR bands was very sensitive to soil optical properties for an erectophile canopy. For a uniform canopy, this band did not produce much significant information for monitoring crop status. Mid-IR provided valuable complementary information on the geometrical structure of the canopy and optical properties of the underlying soil.

INDVI, the time integration of NDVI, is being extensively used to estimate the annual primary production of vegetation over large areas. Goward *et al.*[88] related annual INDVI to NPP of different biomes in North America, and obtained high correlations between INDVI and NPP, but validation of this contribution on continental scales was difficult. Since estimation of NPP from INDVI alone, is not free from lacunae, attempts were made to combine the ecosystem process models, that respond similarly with reference to the criticality with NDVI[89] e.g. as in FOREST−BGC (Bio-Geo Chemical Cycle) ecosystem model. Unlike the man-made agricultural vegetation, the forest productivity is influenced by conditions such as subfreezing temperatures and summer water stress severely constrain gas exchange processes, and yet these were not represented in the NDVI, which responds to leaf area and reflected solar radiation. Thus; integrated NDVI was modified by incorporating water stress related impairment in plant activity through reduction in the effective NDVI at each integration with the surface resistance factor, which was derived from the ratio of

NDVI and surface temperature duly normalized by the slope of NDVI and surface temperature graph (drawn using multiple observations over the forest zone) for climatically similar zones of the scene. Gupta[74] has related INDVI over different 15 days segments of the CGC, starting from the mid of late tillering stage, with the district level wheat yield.

Radiation, Evapotranspiration and Productivity Relations

The growth and decay profile of a crop canopy is expected to be related to crop conditions and hence, to the grain yield. The maximum green biomass attained decides the amount of interception of Photosynthetically Active Radiation (PAR). The energy intercepted at this stage is used mainly for grain formation as there is no more growth in the green biomass, assuming the availability of PAR.

Early studies of Spectral Vegetation Indices (SVIs) indicated the diagnostic nature of canopy green leaf area index as green biomass and per cent green foliage ground cover. However, nonlinearity and variability in LAI relation are their main limitations. More recently, relations between SVIs and Absorbed Photosynthetically Active Radiation (APAR) are being considered as alternative explanation. APAR is more linearly related to SVIs than LAI and has a better direct link to plant photosynthetic processes. But these are also not completely free from nonlinearity and variability. Variations in canopy optical properties and architecture as well as background spectral reflectance also cause deviations in APAR − NDVI relation, with the background spectral reflectance properties producing the largest effect[90].

Despite good correlations between NDVI and canopy features such as biomass, LAI, APAR and canopy photosynthetic capacity, NDVI often fails to capture dynamic physiological processes, which may occur on fine temporal or spatial scales. For example, drought tolerant evergreens can undergo seasonal changes in light-use efficiency and photosynthetic activity without equivalent changes in NDVI or canopy structure[91]. Further, many species render seasonal stress under dramatic mid-day 'photosynthetic depressions'[92] presumably without significant changes in canopy architecture.

Sellers[93,94] derived an important interrelationship amongst LAI, APAR and NDVI that clarifies the utility of these biophysical variables. He observed that APAR was linearly related to NDVI and curvilinearly to LAI approaching asymptotically an LAI of 6, where virtually all incident shortwave radiation (ISR) is absorbed by the canopy. Functionally, APAR is expressed as

$$\text{APAR} = f[\text{LAI, ISR, Canopy Geometry}]$$

The AVHRR NDVI functional relationship with APAR and LAI could be expressed as

$$\text{NDVI} = f[\text{APAR}] = f[\text{LAI}] \qquad \ldots (35)$$

Therefore, for large scale approximations, for a canopy of known structure, light scattering and absorbing properties, any one measure of the canopy can be used interchangeably with the others with some algebraic manipultion of formulae.

Since the general circulation models (GCMs) do not realistically define the land surface biological feedbacks their predictions of global temperature increases are being challenged. Modelling of simple surface vegetation variables is now being incorporated into the latest generation of GCMs. It has become increasingly clear that merely classifying a global vegetation is inadequate and quantitative assessments of energy/mass flux activity and net primary productivity (NPP) are needed to answer critical questions in global ecology and for interface with the GCMs.

The relationship between annual NPP of crops and the integrated APAR absorbed over a growing season is given by

$$\mathrm{NPP} = f[\Sigma\,\mathrm{APAR}]*\epsilon, \qquad \ldots(36)$$

where ϵ is the energy conversion efficiency (g/MJ). Here, the APAR could be functionally replaced by LAI or NDVI.

At times, a subtle but important distinction is drawn between the intercepted PAR (IPAR) and the absorbed PAR since intercepted radiation does not explicitly consider radiation absorption and that not all intercepted photons need to be absorbed, some are scattered (say 15%).

Hence, IPAR is defined as the difference in the PAR flux density above and below the plant canopy. This requires incoming PAR flux density I_0 and the transmitted PAR flux density (T_0) through canopy to the soil surface. Thus, IPAR is given by

$$\mathrm{IPAR} = I_0 - T_c \qquad \ldots(37)$$

From this, capacity of the plant to intercept some fraction (f_i) of the incident PAR flux is estimated as $f_i = \mathrm{IPAR}/I_0$.

The experimental observation by Gallo and Daughtry[95] showed 3.5% difference between PAR and APAR from planting until just before physiological maturity of crop, thus, indicating that IPAR is a reasonable approximation of APAR as long as full green canopies are present.

Although instantaneous differences between APAR and IPAR are usually small, temporal integration of these values, may lead to considerable difference in estimates of PAR available to the plants over the growing season. Absorbed PAR is the algebraic sum of incoming and outgoing flux densities measured above and below a canopy. The two measurements required above the canopy are (*i*) incident PAR flux density I_0 and (*ii*) extant PAR flux density (R_c) from the canopy, including the PAR reflected by the soil, but not absorbed prior to arriving at the top of the canopy while the two measurements required below the canopy are: (*i*) PAR flux density transmitted (T_c) through the canopy to soil surface and (*ii*) PAR flux density reflected by the soil (R_s). From these measurements APAR may be computed as:

$$\mathrm{APAR} = (I_0 + R_s) - (T_c + R_c) \qquad \ldots(38)$$

This APAR is measured in photon units (μmol $\mathrm{m}^{-2}\ \mathrm{s}^{-1}$) or energy units ($\mathrm{Wm}^{-2}$). To make it independent of flux density considerations, the influence of canopy optics and structure on APAR are to be considered separately. This canopy structural or 'state' attribute is the capacity of plant canopy at a given time to absorb some fraction (f_A) of the incident PAR flux, defined as $f_A = \text{APAR}/I_0$.

Remotely sensed observations typically provide only limited temporal samples of plant canopy conditions and in vegetation production models instantaneous APAR is sought since it does not include a rate of PAR accumulation and temporal integration of APAR measurements accomplishes this task through

$$\text{APAR}(t) = \int_{t=0}^{m} f_A(t)\, I_0(t)\,\mathrm{d}t \qquad \ldots(39)$$

Here, integration is done from emergence ($t=0$) to maturity ($t=m$). This seasonally integrated APAR incorporates both the meteorological conditions (incoming solar radiation) and the vegetation variables (leaf area and canopy radiation absorption characteristics). Both variables directly determine the photosynthetic production of a plant. With the functional equivalence as shown in the above equation $\text{NPP} = \Sigma\text{NDVI} * \varepsilon$.

The factors converting annual APAR energy in mega Joules (MJ) to NPP in grams varies with biome type e.g., for crops it ranges from 1 to 3 g/MJ; for grass lands 0.8 and deciduous trees 1.0 etc. Here, a static NDVI is not used, rather the annual integration of weekly composited NDVI, analogous to the integrated APAR is applied. Seasonal interpretation of the weekly composites automatically reduces the cloud contamination and improper climatic sensitivity of the relationship. Although the maximum NDVI or LAI of many biomes in many climates is rather similar, higher productivity regions have substantially larger periods of active vegetation growth, reflected in larger value for integrated NDVI.

The general assumption for INDVI are more true for crops, but lesser for perennial vegetation such as forests or greenlands. This vegetation is routinely water stressed, usually temperature limited at some point of time each year, rarely has optimum nutrient availability and often fixes substantial amounts of carbon during dormant periods while not visibly growing. In the translation from NDVI to NPP, much of this variability is normally subsumed in the ε coefficient, which then incorporates both meteorological and biochemical components.

Running[96] developed a more explicit formulation that separates meteorological from biochemical components, so that it should more generally be applicable under a wide variety of conditions and biome types.

$$\text{NPP} = t\Sigma T(\sigma + \text{NDVI}) * \varepsilon, \qquad \ldots(40)$$

where σ = Surface resistance = T_s/NDVI

$t\Sigma T$ = Temperature controlled NDVI time integration

T_s = Surface temperature defined by AVHRR thermal IR split channels 4 and 5 (10.3-12.5μm atmospheric window is split into 10.3-11.3 and 11.5-12.5μm windows to take advantage of differential absorption by water vapour in these two spectral regions).

Because only the standard AVHRR channel data are required, ultimately this formulation could be implemented globally just as in the current global NDVI maps. Hope *et al.*[97] and Nemani and Running[89] have suggested that variations in the slope are diagnostic of differences in latent heat of evaporation fluxes or surface parameters affecting these fluxes.

Spectral vegetation indices based on satellite data have been used to assess the spatial pattern in vegetation condition in terms of above ground green biomass, LAI (e.g., Running *et al.*[98]) or crop yield (e.g., Barnett and Thompson[99]). Annual or growing seasons' integrates of NDVI and other indices have been related to annual net primary production.

The success of vegetation indices to illustrate spatial and temporal variability in above ground vegetation arises from their relationship to intercepted PAR, canopy photosynthetic capacity, and maximum canopy conductance. Since maximum canopy conductance constrains surface H_2O and CO_2 fluxes, NDVI may prove valuable as a forcing variable in biosphere-atmosphere exchange models linked to atmospheric circulation models.

Since vegetation is considered to be the functional equivalent of terrestrial ecosystem, vegetation serves as prime indicator for ecosystems such as grasslands, forests and deserts. The characterization of vegetation generally offers the observer a myriad of information about its geographical location, current ecological conditions and previous changes that might have occurred in the region.

In the land surface-atmosphere interactions, evapotranspiration is, an important process. Development of biophysically-based vegetation models and their incorporation into General Circulation Models (GCMs) greatly improved the realisation of land surface parameterization and further efforts are now being put in that direction. The three basic land surface properties known to be ultimately governing these interactions are the albedo (in radiative transfer), surface roughness (in momentum transfer) and surface hydrological parameterization (in sensible and latent heat exchange via land and vegetation). Till the last decade, most studies dealt these three properties independently. The sensitivity studies have now shown that these need to be dealt inter-dependently for explaining the processes involved in the GCMs. Solar radiation provides almost all of the energy received at the surface of the earth; while some of the radiant energy is reflected back to space, a part is reradiated in the thermal IR waveband. The net radiation, i.e., energy remaining at the earth's surface, drives physical processes. The energy balance may be expressed as:

$$R_n = H + LE + G + P_s \qquad \ldots (41)$$

Here, R_n, H, LE, G and P_s refer to net radiation, sensible heat flux, latent heat flux, ground heat flux and energy used for photosynthesis, respectively.

The significance of this equation is that it is applicable at the scale of a single plant as well as cropped field, explaining how energy is provided to warm the soil and crop and to evaporate water. More important is that this equation is no less valid on the global scale, explaining how energy is provided to the continents and oceans, where vast quantities of heat and vapour are given to or extracted from the atmosphere.

To compute evapotranspiration using remote sensing approach, the sensible heat flux (H) of eq. (41) is defined in terms of the temperature difference between the surface of the vegetation canopy (T_s) and the ambient air (T_a). Thus, the surface temperature "T_s" is related to latent heat flux LE by the well known energy balance equation during the day time conditions.

$$LE = R_n - G - \rho(C_p/r_a)(T_s - T_a). \qquad \ldots (42)$$

Here, ρ, C_p and r_a refer to air density, specific heat of air at constant pressure and aerodynamic resistance, respectively.

This equation is the most frequently used one for crop water status discrimination on a local scale or for surface fluxes estimation on a regional scale. The main application of this equation lies in its capability of giving flux measurement methods for soil-plant-atmosphere modeling. For the long term monitoring of crop water conditions, periods between 1 to 10 days are to be considered, especially for adjusting to the time scale of agrometeorological processes and related models.

For example, the taller crop/vegetation with aerodynamically more rough nature extracts more amounts of energy from the ambient air. Similarly, broad leaved plants transpire more than that of the grasses for the reason that larger beings tend to be less efficient in dissipating through convective transfer and will therefore have more energy available for evapotranspiration than the plants with small and narrow leaves. In advection, if latent heat is more than $R_r + H$, then sensible heat will be drawn from the air and consumed in evapotranspiration. Albedo also varies with the vegetation vigour as influenced by the greenness and the relative water content (turgidity). Thus, these three properties in turn constitute the important and mutually dependent land surface parameters to be accounted for in dealing with the GCMs.

Plants need large quantities of water, e.g., about 1000kg water is required to produce 1kg of wheat, as actively growing plant uses 5-10 times as much water, every day, as it can hold at one time. On the other hand, if plants are exposed to nearly 100% relative humidity, a reduction in growth is observed because of the interruption in translocation of the plant products. Transpiration is highly involved in the transport of essential nutrients from soil and plant tissue where the nutrients enter into the photosynthetic process.

When the rate of water intake to plants cannot equal the rate of water loss, stomatas either close partially or completely, thereby increasing stomatal resistance, which is an important parameter in the computation of ET. In the energy balance equation; ET is major consumer while R_n is major energy supplier, except in special circumstances when H is also a significant source of energy. In humid regions, R_n is major source of energy for ET, since R_n puts an upper limit on the energy consumed as ET. In aird regions, the advection of sensible heat from warm, dry regions provides an important additional energy source. Differences in cultural practices or morphological differences in plants alter the energy balance over plant canopies. When albedo of a vegetation surface is increased, R_n is decreased and ET rates will be reduced.

As per Penmann[100], the potential Evapotranspiration (ET_p) is defined as the ET rate of the "short green crop, completely shading the ground, of uniform height and never short of water". It actually measures evaporative power of the atmosphere. The term potential is being changed to reference, to relate the concept more directly to crop water requirements, specifying the reference crop (usually short grass or *alfa alfa*) and assuming that the leaf surfaces are typically not wet.

ET_p is determined by climatic factors, including temperature, radiation, humidity and wind speed. Actual ET can differ from potential ET in situations where soil moisture conditions are limiting, or where vegetation characteristics differ from those specified in the definition of ET_p. The practiced-in method for computing ET_p over crops depends on the data availability as given in Table IV[100-108] for various methods given in Table V[109]. Temperature is an indicator of the evaporative power of the atmosphere and this forms a primary input in each method. When measured values of R_n are not available, these could be computed, using solar radiation, maximum and minimum values of temperature and humidity (Allen *et al.*[110]). To estimate crop ET, potential or 'reference' ET is first computed, then appropriate crop coefficients are applied to reflect seasonal and crop differences in water use. The only method which is capable of reflecting actual ET rates under conditions of limited soil moisture is the Penman-Monteith equation (Table V).

ET could also be computed using water balance methods wherein the terms in the water balance equation could be deduced on hourly basis from micrometeorological measurements, by either energy balance or eddy-correlation approach. By accounting for the precipitation, infiltration, deep percolation and run off/seepage over the soil surface in a unit time, the changes in soil moisture could be accounted for by the evaporation from the soil surface and the uptake of crop plants.

Table IV

Data requirements of selected formulae for computing potential evapotranspiration

Method	T	R_s	e	u	Latitude/ day-length	Plant Parameters	Temporal resolution of input data
Thornthwaite	X				X		monthly
Blaney-Criddle (SCS)	X				X		monthly
Hargreaves	X		X		X		monthly
Samani-Hargreaves	X				X		daily[a]
Jensen-Haise	X	X					daily
Priestly-Taylor	X	X					daily[a]
Penman	X	X	X	X			daily[a,b]
Penman-Monteith	X	X	X	X		X	daily[a,b]

T, temperature; R_s, solar radiation; e, humidity; u, wind.

[a] Requires daily minimum, maximum temperature.

[b] Can be used with hourly input data.

Table V

Formulae for eight methods of computing potential or 'reference' ET, expressed in a form which emphasizes differences in data requirements and computational structure

Method	Formula
Thornthwaite	$\mathrm{ET}_p = c \star f(\mathrm{LAT}, \mathrm{JD})m \star [T_m/f_1(T_m)]f[F_1(T_m)]$
Blaney-Criddle (SCS)	$\mathrm{ET}_p = [(c_1 \star T_m) - c_2] \star [T_m \star f(\mathrm{LAT}, \mathrm{JD})m]$
Hargreaves	$\mathrm{ET}_p = f(\mathrm{LAT}, \mathrm{JD})m \star T_m \star f(T_m, e_m)$
Samani-Hargreaves	$\mathrm{ET}p = c \star R_t \star T \star (T_{\max} - T_{\min})1/2$
Jensen-Haise	$\mathrm{ET}p = R_s \star f(T, \mathrm{ELEV}) \star (T - f(T, \mathrm{ELEV}))$
Penman	$\mathrm{ET}_p = \dfrac{f_1(T) \star R_n + f_2(T) \star f(u) \star [f_3(T) - f(T,e)]}{f_1(T) + f_2(T)}$
Priestly-Taylor	$\mathrm{ET}_p = c \star \dfrac{f_1(T) \star R_n}{f_1(T) + f_2(T)}$
Penman-Monteith	$\mathrm{ET}_p = \dfrac{f_1(T) \star R_n + f_1(T) \star f_2(T) \star [f_3(T) - f(T,e)]/(u, HT)}{f_1(T) + f_4(T) \star [1 + r_c/f_1(u, HT)]}$

C, constant term; *T*, temperature; *f* (variable), term computed as a function of (variable); R_t, extra-terrestrial radiation; R_s, solar radiation; R_n, net radiation; *e*, humidity (actual vapor pressure); *u*, windspeed; r_c, canopy resistance; LAT, latitude; HT, vegetation height; ELEV, elevation; JD, julian day.

Subscripts: *m*, monthly value of variable; (no subscript), assumed daily value of variable; 1,2, etc., used to denote different functions of the same variable(s) within an equation, or different constants; max or min, daily maximum or minimum value of variable.

In energy balance method, the Bowen ratio (β) of the sensible heat flux H to the latent heat flux LE is calculated as

$$\beta = H/LE = \gamma \delta T/\delta e, \qquad \ldots (43)$$

where γ is psychrometric constant, δT and δe are differences in air temperature and water vapour pressure between two levels above the canopy. Latent heat flux is then computed from the energy balance equation:

$$LE = (R_n - G)/(1 + \beta), \qquad \ldots (44)$$

where R_n is net radiation, and G, the heat flux into ground.

In case of forests, the aerodynamically rough surface, δT and δe are small and thus difficult to measure. Besides this, the layer just above a forest is influenced by individual surface elements (the trees) and the exact similarity between heat and water vapour in this layer is in question. Thus the eddy correlation method is often preferred to estimate ET above forests. It is based on the fact that on the average, air rising above forest carries extra humidity coming from evaporation whereas air going down is drier. Here, evaporation (E) is computed using

$$E = \rho_e' W', \qquad \ldots (45)$$

where ρ'_e is the fluctuation in water vapour density around its average value and W' is the fluctuation in the vertical component of wind speed. This formula has the advantage of making no assumption on the turbulent transfer of either heat or water vapour. Here, ρ'_e is measured with a fast response hygrometer (time constant $W = 0.5$ sec) such as an ultraviolet gas analyzer; W' is measured with devices having very less distant constant such as sonic anemometer.

Once E gets known, canopy resistance could be directly calculated if leaf temperature is measured or computed individually using Penman-Monteith equation, i.e.,

$$LE_0 = [\Delta(R_n - G) + (\rho C_p / r_a)\delta e] / [\Delta + \gamma(1 + r_c / r_a)], \qquad \ldots (46)$$

Δ = Derivative of function relating Saturated Vapour Pressure to temperature,
δe = water vapour saturation deficit,
r_c = canopy resistance
and
r_a = aerodynamic resistance.

The aerodynamic resistance is computed using

$$r_a = [\ln (z_0 - d)^2 / z_0] / (k^2 u), \qquad \ldots (47)$$

where z_0, d, k, and u refer to roughness length (m), zero plane displacement height (m), von Karman's constant (0.41) and wind speed at 2m height (ms^{-1}), respectively.

Forests are characterized by small values of r_a of $5\text{-}10 sm^{-1}$ but with larger values of r_s w $100sm^{-1}$, and thus the equation (46) could be reduced to

$$LE_c = \rho C_p (\delta e / r_c) \qquad \ldots (48)$$

r_c may be estimated from L (latent heat) and stomatal resistance $r_s \cdot r_c$ can be taken as r_s/L where r_s is average stomatal conductance. Stomatal resistance depends on light, temperature, vapour pressure deficit and soil moisture.

In studying the regional crop or natural vegetation conditions, thermal infrared domain offers surface temperature, which could be obtained from geostationary satellites like INSAT, METEOSAT and GOES or near polar orbiting meteorological satellites like NOAA on a regular basis. NOAA/AVHRR data are available since more than thirteen years with spatial resolutions of 1.1km (LAC) or 4km (GAC). The availability of two thermal infrared channels from NOAA, allowing to use the near linear split-window combinations, ensure a precision of the order of about 2 °K. Landsat TM channel 6 (10.5-12.5μm), with a spatial resolution of 120m and a repetitivity of 16 days also offers possibilities for precise studies of local microclimates on selected dates.

The surface temperature derived is indirectly related to the evaporation flux in the energy balance equation. Monteith[111] and Jackson[112] have well established the physical meaning attached to the surface temperature through studies conducted over a large number of field studies at a local scale based upon detailed energy balance and micrometeorological measurements. These arguments have stimulated a large number of studies for the estimation of the latent heat flux using soil-plant-atmosphere modeling.

For larger scales, Seguin[113] developed an idea, originally proposed by Jackson *et al.*[114], who related daily evapotranspiration ET_d to accumulated daily net radiation R_{nd} and mid day value of $(T_s - T_a)$. The resulting linear equation obtained by this simplified approach is:

$$ET_d = R_{nd} + a - b(T_s - T_a) \qquad \ldots (49)$$

with '*a*' and '*b*' being adjustment constants, respectively, expressed in mm/day or mm/day/°C. These are mainly a function of surface wind velocity and stability factor. Seguin and Itier[115] gave conditions for passing from instantaneous energy balance equation to the above equation.

When this relation was applied to time scales larger than the daily basis, dispersion for individual leaves (of the order of ±1 mm/day) is lumped by passing to 5 or 10 days. The accumulation of daily values over *n* days gives

$$ET_d = \sum R_{nd} + na - b \sum (T_s - T_a). \qquad \ldots (50)$$

Computation of kinetic temperature from radiometric temperature needs information on emissivity. Schmugge *et al.*[116] have discussed the variations in emissivity in thermal IR region using HAPEX data sets. The emissivity could be related to NDVI[117]. Sellers *et al.*[118] have presented the interim results of the First ISLSCP (International Satellite Land Surface Climatology Project) Field Experiment (FIFE) wherein the method of computing *ET* with remote sensing inputs has also been presented.

Spatial Integration and Global Circulation Models

Realistic modeling of land surface processes is essential for successful simulation of climate, but there are many obstacles. To improve land surface hydrological models for GCMs we must now learn how to combine effects of processes from local to GCM grid scales in a rigorous manner wherein space observation have an immense potential to give the inputs.

For the monitoring of dynamic vegetation resources over earth, especially for understanding global change processes, AVHRR based NDVI is assuming predominant role. The interrelationships between high spatial (30 to 80m) and low spatial (1.1km) resolutions NDVIs have been studied through spatial integration processes[119-122]. Based on the spatial analysis by Fourier and scale invariance analysis, Townshend and Justice[123] have recommended 250m spatial resolution for land transformation studies. The RVI and NDVI are derivable from monitoring (NOAA/AVHRR) as well as from resources management (Landsat, IRS, SPOT etc.) satellites.

High spatial resolution data are often used for simulation purposes and are an excellent tool in preparing for future space missions. In investigating the spatial integration of NDVI, methods based on averaging, weighted averaging, K-near, median etc. could be employed for smoothening the frequency distribution and reducing the noise. Aman *et al.*[119] spatially averaged the data for spatial degradation from original to coarser data. For a spatial resolution lower than 1000m (300-1000m), the integrated NDVI, INDVI, was found to be linearly correlated to the spatial average of NDVI, when the study was conducted over two West African tropical sites and one French temperate site. Raffy[124] has discussed the general frame for the development of models when scale of input measurements is changed.

To investigate and resolve structural and spectral variations within the vegetated land scales, Goward and Huemmrich[84] suggested bidirectional measurements and increased spectral coverage. But both these approaches involve considerable investment in sensor design and operation as well as processing requirements.

Before opting for any of these approaches, the complexity involved in solving the problem needs to be demonstrated. Goward and Huemmrich[84] themselves suggested to undertake short-term "earth-probe" space missions such as AVHRR sensors in geostationary orbit to provide data at multiple sun angles and a simple silicon grating spectrometer in polar orbit for multispectral observations on a pilot basis to address the structural variability and spectral attributes of vegetation communities across the globe.

To a reasonable extent NDVI, though it is not a linear combination of reflectances, can be spatially integrated without significant loss of information. In addressing problems linked to spatial integration and scale change, use of remotely sensed data in estimating the biophysical parameters of vegetation by dynamic models would be of much utility. NDVI, being integrable, linearly coupled its significant relation to the canopy parameters is a more useful radiometric index for global parameterization of vegetation related processes such as primary production and heat and mass transfer between the land surface and the atmosphere.

The remote sensing community has steadily gained more knowledge about the scattering behaviour of vegetation canopies. Till now the community has relied only on nadir data and statistical pattern recognition techniques to extract surface properties. Kimes *et al.*[125] believed that for significant advances in the accuracy of remotely sensed references, one must utilize additional knowledge. They developed a knowledge based expert system (VEG) for inferring physical and biological surface properties. VEG accepts nadir and/or bidirectional spectral data of an unknown target as input, determines the best strategy or strategies for inferring physical and biological surface properties, applies the strategy or strategies to the target data and provides a rigorous estimate of the accuracy for the inference. The VEG and the anticipated system expansions would provide a valuable tool for intense study of the sites using spectral data from the current as well as the forthcoming sensors on EOS. It is expected in future to develop it further for accounting the atmospheric effect and predicting any nadir or off-nadir reflectance given any combination of nadir and/or off-nadir reflectances of a target.

Vegetated surfaces complicate the problem of representation of the land surface, for its influence on the sensible and latent heat fluxes. Although the need for GCMs to include vegetated surfaces is well accepted, attempts have so far been limited. Until recently, the land surface properties which regulate the exchange of radiation, momentum and heat with the atmosphere have been regarded as separate item which may be independently prescribed as boundary conditions within a GCM of the atmosphere. The three land surface properties which ultimately govern these interactions are the albedo (radiation transfer), surface roughness (momentum transfer) and the surface hydrological parameterization/prescription (sensible and latent heat transfer). These parameters when subjected to sensitivity analysis were found to have important impacts on atmospheric fields. Dickinson[126] pointed out that these experiments were unrealistic particularly with regard to the description of land surface evaporation, as specified in the GCMs. Therefore, Dickinson[126] designed a model which explicitly refers to the effects of vegetation. Sellers *et al.*[127] constructed a Simple Biosphere (SiB) model which incorporates much of the philosophy of the Dickinson[126] model while differing in detail. This is based on the 'big leaf' concept, in which the vegetation is represented by a layer of negligible heat capacity. In grid cells with incomplete vegetation cover, the model assigns a proportion of the cell to bare soil and others and remaining proportion to vegetation; thus, the determination of surface energy flux of sensible and latent heat includes contributions from both bare soil and vegetation canopy. SiB has eight prognostic physical state variables—three temperatures (canopy-vegetation, T_c; Soil cover, T_{gs}; deep soil temperature T_d); two interception water stresses (one for canopy M_c, and one for ground cover, M_g) and three levels of soil moisture storages (W_1, W_2 and W_3). Time dependencies of T_c and T_{gs} are described by energy balance equation which includes radiation, sensible heat, latent heat and heat storage term. The parameters in radiation calculation include LAI, leaf angle distribution functions, leaf optical properties (live and dead) averaged for visible, near-IR and thermal IR wave length intervals, green leaf fraction and canopy cover fraction. There are three surface resistances: the bulk canopy resistance r_c, the ground cover resistance r_g and the soil surface resistance r_s. SiB was originally designed to be as realistic as practicable with a partial view towards utilizing satellite data for its initialization and validation. Studies continue in this area with particular emphasis on using satellite data in the visible, near IR and thermal IR wavelength intervals.

Global Change and Future Earth Observation Missions

Vegetation is an .important parameter for global change studies. Satellite remote sensing offers a highly feasible means for collecting appropriate, consistent and temporally frequent data over the entire geosphere-biosphere, giving representative information of the processes and interactions responsible for global change, that too at a reasonable cost.

The International Geosphere-Biosphere Programme (IGBP) is the current on-going activity for global change studies and consists of core projects namely the Biological Aspects of Hydrologic Cycling (BAHC), Global Change and Terrestrial Ecosystem (GCTE), International Global & Atmospheric Chemistry (IGAC), Joint Global Ocean Flux Study (JGOFS), Land Ocean Interactions in Coastal Zones (LOICZ) and Past Global Changes (PAGES)[122,128,129]. IGBP also has over-arching and integrative activities mainly Global Analysis, Interpretation and Modeling (GAIM), a Data and Information System (IGBP-DIS), and a System for Analysis, Research and Training (START). With a view to under-

Table VI

Details of US NASA earth observing system (EOS) programme (NASA, 1993)

Launch	Spacecraft	Lifetime (Years)	Instrument Complement					
1998	AM1	5	MODIS	MISR	CERES (2)	MOPITT	ASTER	
2003	AM2	5	MODIS	MISR	CERES	EOSP	TES	MOPITT*
2008	AM3	5	MODIS	MISR	CERES	EOSP	TES	
1998	COLOR	3	SeaWiFS-Type					
2000	AERO1	3	SAGE III					
2003	AERO2	3	SAGE III					
2006	AERO3	3	SAGE III					
2009	AERO4	3	SAGE III					
2012	AERO5	3	SAGE III					
2000	PM1	5	MODIS	AMSU	MIMR	AIRS	MHS	CERES (2)
2005	PM2	5	MODIS	AMSU	MIMR	AIRS	MHS	CERES
2010	PM3	5	MODIS	AMSU	MIMR	AIRS	MHS	CERES
2002	ALT1	5	GLAS	TMR	SSALT	DORIS		
2007	ALT2	5	GLAS	TMR	SSALT	DORIS		
2012	ALT3	5	GLAS	TMR	SSALT	DORIS		
2002	CHEM1	5	HIRDLS	SOLSTICE II	ACRIM	MLS	SAGE III	TBD(J)
2007	CHEM2	5	HIRDLS	SOLSTICE II	ACRIM	MLS	SAGE III	TBD(J)
2012	CHEM3	5	HIRDLS	SOLSTICE II	ACRIM	MLS	SAGE III	TBD(J)

Notes:
1. CERES(2)—means two scanners
2. AIRS, AMSU, HIRDLS, MHS, MIMR and MODIS data available via direct broadcast
3. ASTER data available *via* direct downlink
4. CERES and LIS funded for TRMM-1 (Tropical Rainfall Mapping Mission) in 1997
5. NSCAT II planned for flight on ADEOS-II (Advanced Earth Observations Satellites of Japan)
6. EOS – AERO Spacecraft to be provided through international cooperation
7. TBD(J): Japanese Instrument to be Determined

*Accommodation of MOPITT on EOS – AM2 to be confirmed.

Note: Please see list of acronyms on next page.

List of acronyms used

ACRIM : Active Cavity Radiometer Irradiance Monitor
ADEOS : Advanced Earth Observations System (from Japan)
AIRS : Atmospheric Infrared Sounder
AMSU : Advanced Microwave Sounding Unit
ASTER : Advanced Spaceborne Thermal Emission and Radiometer (provided by Japan)
AVHRR : Advanced Very High Resolution Radiometer
CERES : Clouds and Earth's Radiant Energy System
DORIS : Doppler Orbitography and Radiopositioning Integrated by Satellite (provided by France)
EOSP : Earth Observing Scanning Polarimeter
GLAS : Geoscience Laser Altimeter System
HIRDLS : High Resolution Dynamic Limb Sounder
JERS : Japan's Earth Resources Satellite
MHS : Microwave Humidity Sounder (Provided by European Organisation for the exploitation of Meteorological Satellites)
MIMR : Multifrequency Imaging Microwave Radiometer (provided by European Space Agency)
MISR : Multiangle Imaging Spectroradiometer
MLS : Microwave Limb Sounder
MODIS : Moderate-Resolution Imaging Spectroradiometer
MOPITT : Measurement of Pollution in the Troposphere (Canadian)
NOAA : National Oceanic and Atmospheric Administration
POLDER : Polarisation and Directionality of the Earth's Reflectance
SAR : Synthetic Aperture Radar
SAGE : Stratospheric Aerosol and Gas Experiment
SOLSTICE: Solar Stellar Irradiance Comparison Experiment
SSALT : Solid State Altimeter (provided by France)
TES : Tropospheric Emission Spectrometer
TMR : TOPEX Microwave Radiometer

stand processes of Earth systems science US-NASA has designed Earth Observations System (EOS) programme during 1998-2015 A.D. and Table VI gives the details of various categories of programme elements and on-board sensors[130]. The EOS programme will consist of AM, COLOR, AERO, PM, ALT and CHEM series catering to descending node 1030hrs equator crossing earth and atmosphere remote sensing, ocean colour, aerosol, ascending node 1330hrs equator crossing earth and atmosphere remote sensing, altimetry and chemistry (1530hrs equator crossing time) observations, respectively. EOS-AM, PM and CHEM spacecraft orbit will have 98.2° inclination, 705km altitude, 16 day/233 orbit repeat cycles. The EOS-ALT and EOS-AERO would have sun-synchronous polar and 57° inclined orbits, respectively. To complement the NASA-EOS programme with resource management- and monitoring-satellites, the European Space Agency (ESA) shall be launching POEM-ENVISAT (Polar Orbit Earth Observation Mission-Environmental Satellite) from 1998 A.D. onwards and POEM-METOP (Meteorological Operational Satellites) from 2005 A.D. onwards while Japan would cater through ADEOS

(Advanced Earth Observation Systems) from 1999 A.D. onwards. Data from NOAA polar orbiting meteorological satellites would continue to be available till 2002 A.D.

The terrestrial surface is linked to atmosphere through fluxes of energy and trace gases *via* vegetation and other processes. The vegetation plays the role of partitioning the incoming solar energy and as a consumption function for evapotranspiration. This with soil as a storage function and climate as forcing function define the conceptual needs of BAHC and requires measurements of meteorological variables, vegetation, topography and snow cover parameters. Study of consequences of global change on structure and function of ecosystems, and prediction of potential feedbacks of the terrestrial changes on atmosphric, climatic and landuse changes as visualized in GCTE would need information on leaf area, canopy architecture, net primary productivity, trace gases efflux, air temperature above planetary boundary layer, canopy surface temperature, soil temperature, soil moisture etc.

The forthcoming sensors on EOS would provide very enhanced capabilities to study agriculture and natural vegetation together with possibility of incorporating atmospheric and view angle corrections based on measurements. Moderate-Resolution Imaging Spectroradiometer (MODIS), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Multi-Angle Imaging Spectroradiometer (MISR), Multi-frequency Imaging Microwave Radiometer (MIMR) and Tropospheric Emission Spectrometer (TES) would cater to these requirements, whose salient features are discussed as under.

MODIS is a primary and truly multidisciplinary instrument on EOS-AM as well as EOS-PM and caters to Earth processes studies over land surface, ocean and atmosphere[131]. The measurements would be made in 20 bands between 0.4 and 3.0μm and 16 bands between 3 and 15μm. It will have spatial resolution of 250m (for 2 bands), 500m (for 5 bands) and 1000m (for 29 bands), signal-to-noise ratio greater than 500 at 1-km resolution (at a solar zenith angle of 70°); absolute irradiance accuracy of $\pm 5\%$ from 0.4 to 3μm (2% relative to sun) and 1% or better in the thermal infrared (3 to 15μm), surface temperature resolution of 1°K over land and 0.2°K over oceans, polarization sensitivity of less than 2% for the visible out to 2.2μm, and would provide continuous global coverage every 1 or 2 days. In case of vegetation/land surface cover it would provide vegetation indices corrected for atmosphere, soil and directional effects; estimation of NPP, LAI, IPAR; and information on land cover type with capabilities for change detection and identification. It will have bands to compute optical thickness, particle size and mass transport of aerosols; and global distribution of total precipitable water.

ASTER would be the Japanese instrument on board EOS-AM1 and will operate in three visible and near-infrared (VNIR) bands between 0.5 and 0.9μm with 15m resolution; six shortwave infrared (SWIR) bands between 1.6 and 2.5μm, with 30m resolution; and five thermal infrared (TIR) bands between 8 and 12μm, with 90m resolution. The swath width would be 60km at nadir and the swath center is pointable cross-track $\pm 8.5°$ in SWIR and TIR, with the VNIR pointable out to $\pm 24°$. This would enable accessibility to any point on the globe at least once every 16 days in all the 14 bands and once every 5 days in the three VNIR bands. The combinations of VNIR and SWIR bands would provide information on vegetation and land use while TIR bands would provide information on surface kinetic temperature and spectral emissivity.

The High Resolution Imaging Radiometer (HIRIS) is likely to be available on EOS-AM2, probably in place of ASTER. The instrument would measure reflected radiance in 0.40 to 2.5μm wavelength region in 192 spectral bands with contiguous samples of approximate 10nm width[132]. The spatial resolution at nadir will be 30m and swath width would be 24km. The 10nm spectral sampling in visible and near IR region permits biological studies of vegetation while its availability in short wave IR is useful in mineralogical mapping. The $\pm 45°$ cross-track pointable capability would permit sampling of a given point on the earth's surface at a minimum of every 2 days. The along-track pointing capability of $+56°$ to $-30°$ (positive sign refers to velocity direction) would be useful to estimate the bidirectional reflectance distribution function of surfaces, to remove atmospheric attenuation, and to implement image-motion compensation to increase signal-to-noise ratio for dark targets.

MISR, on-board EOS-AM series, would provide nine discrete views of the object in four spectral bands centered at 0.443, 0.555, 0.67 and 0.865μm. The nine views would consist of nadir and four other symmetrical for-aft views upto $\pm 70.5°$ forward and aftward of nadir. Each of these 36 data channels (four spectral bands for each of the nine cameras) would be individually commandable to provide ground sampling of 240m, 480m, 960m or 1.92km over the swath width of 356km resulting into coverage of the entire earth once in 9 days at equator and once in 2 days at the poles. The measured angular signatures could be related to canopy structural parameters, which would improve vegetation cover classifications and enable computation of surface hemispherical albedos. Using these information, the APAR, photosynthesis and transpiration rates for the vegetation canopy can be derived.

MIMR, on-board EOS-PM; using six microwave frequencies would provide measurements at 4.86km (90GHz), 11.62km (36.5GHz), 22.3km (23.8GHz), 22.3km (18.7GHz), 38.6km (10.65GHz) and 60.3km (6.8GHz) spatial resolutions. These observations would complement visible, infrared and active microwave observations of vegetation status, biomass and soil moisture. The MIMR will provide data on atmospheric water content and precipitation, to be interpreted in combination with the information on temperature and humidity profiles obtained using Advanced Microwave Sounding Unit (AMSU) and Microwave Humidity Sounder (MHS) data.

TES, on-board EOS-AM, a high resolution infrared imaging Fourier transform spectrometer with spectral coverage of 2.3 to 15.4μm at a spectral resolution of 0.025cm^{-1} would provide information on natural resources of trace gases such as methane from organic decay, including vegetation.

The global aerosol distribution and optical thickness in the troposphere and stratosphere, computable using Earth Observing Scanning Polarimeter (EOSP, on-board EOS-AM) based radiance and polarization measurements in 12 bands in 0.41-2.25μm spectral region, would provide inputs for atmospheric corrections for the study of vegetation and land surface characteristics. The

EOS-AERO mission would further enhance the inforamtion on various types of aerosols at global spatial scale.

Conclusions

Despite current availability of remote sensing data, many of the measurements are only loosely coordinated in time and space. The international collaboration to maintain and coordinate the collection of pertinent Earth System data for global change studies, becoming available from evolutionary Earth Observing System (EOS), is needed.

The climate and land surface subsystems are dynamically coupled through the physical processes of energy and water supply transformation and transport at the land-atmospheric interface. The biosphere interacts with these physical processes to further modify this coupling. Large scale patterns of climate are influenced by regional variabilities of the surface processes of the water cycle. Present physically based models need development which reflect the dynamic coupling in the climate-soil-vegetation system.

In the Soil-Vegetation-Atmosphere-Continuum, certain surface meteorological conditions can be directly calculated, from existing and improved optical satellite sensors, including albedo, short and longwave radiation balance, surface temperature etc. Current satellite NDVI products provide a good measure of seasonal vegetation phenology, but need calibration for viewing conditions and differing bi-directional reflectance properties of different biome types. The current research need is to incorporate the variables as directly measurable by the satellite into the SVAT (Soil-Vegetation-Atmosphere-Transfer) models. Vegetation canopy partitioning into interception (and evapotranspiration) or throughfall and soil storage could be modeled using satellite vegetation indices. For example, Evapotranspiration (*ET*) could be included using the satellite vegetation indices with the appropriate surface meteorological conditions (incident solar radiation, water vapour pressure deficit) to drive an evapotranspiration equation such as the Penman-Monteith equation. Recent studies have suggested that a satellite derived surface resistance factor may also be entered into the equation of *ET*. These simple satellite driven models must be validated with 'point' measurement and mechanistic plant-atmosphere models.

One of the most important means of measuring terrestrial ecosystem dynamics will be through remote sensing, and the data derived from remote sensing becomes an important input necessary for GIS (Geographic Information System) data base. Maintenance of temporal continuity and spatial resolution of remote sensing data sets, long term calibrations of the instrument plus the ground based validation of remote sensing data would aid in long-term interpretation of ecosystem changes.

Remotely sensed data can be used as a tool to detect, monitor and evaluate changes in ecosystem to develop management strategies for ecosystem resources. The basis of change detection is the comparison of remotely sensed and map data or the comparison of remotely sensed data taken at 2 or more different times. The various techniques include transparency compositing, image differencing, use of mixed spectral temporal/spatial/multi-sensor data sets based classified images, image ratioing, principal component analysis, albedo difference images etc.

The existing 1.1km spatial resolution NOAA/AVHRR data based computation of Vegetation Indices and *ET* provide important path finder data sets for monitoring the planet Earth while the earth resources satellites such as IRS, Landsat, SPOT, ERS (European Remote Sensing satellite with active microwave sensors) provide selective diagnostic structural insights within 1.1km AVHRR resolution besides serving to resource management interests. Global, Continental and Regional scale long-term prediction oriented goals would need information on processes and EOS data in 1998-2015 A.D. time frame would make remote sensing better deterministic.

The review has revealed that very little research work has been carried out in India for tapping the full potential of the satellite remote sensing for the studies on vegetation. For example, experimental studies similar to that of SAMMA, and studies dealing with the spatial integration *vis-a-vis*, global circulation models need to be carried out more extensively for understanding the advantages and limitation of the remote sensing technology in a fuller way.

Acknowledgements

Authors are thankful to their organisation, National Remote Sensing Agency, Hyderabad which provided them the opportunity to learn and practice the remote sensing technology which certainly enabled them to write this paper. The technical help of Dr M V R Sesha Sai, Sri S Prasad and Sri T S Viswanadham and logistics support of Sri V V S B Sastri and Ms N Ch Tirumala Devi is acknowledged with gratitudes.

References

1 R B MacDonald *IEEE Trans Geo Sci Rem Sens* **22** (1984) 473

2 S R Schneider, D F McGinnis and J A Gatlin *Use of NOAA/AVHRR Visible and Near Infrared Data for Land Remote Sensing NOAA Tech Rep* NESS **84** Washington DC USA (1981) 44

3 R K Gupta, T S V Nadham, S Prasad and B M Singh *Interrelationships among* 0930, 1030, 1330, 1430*hr Landsat TM and IRS LISS spectral bands NDVI over the wheat growth cycle Preprint IAF Symp* Washington DC (1992) 9

4 R M Hoffer *Biological and Physical Considerations in Applying Computer-Aided Analysis Technique through Remote Sensor Data Remote Sensing—A Quantitative Approach* McGraw Hill New York (1978)

5 M W Gausman *Photogramm Engng* **40** (1974) 183

6 EB Kniplling *Remote Sensing Environ* **1** (1970) 155

7 G Guyot In: *Applications of Remote Sensing in Agriculture* (M D Steven and J A Clark) Butterworths London (1990) 19 *Eds*

8 S G Ungar and S N Goward *Adv Space Res* **2** (1983) 291

9 M C Porwal and P S Roy *Photonirvachak* **20** (1992) 21

10 D N Pant, K K Das and P S Roy *Photonirvachak* **20**(1992) 9

11 S Sudhakar, N Krishnan, Pravin K Das, Atanu K Raha and Arun Gost *ITC Jl* (1992-93) 285

12 R N Jadhav, K Murthi, K K Chouhan, M N Tandon, M L Arravatia, V K Srivastava, A K Kanda, G V S Baba, M P Oza, M M Kimothi, K Sharma and J K Thesia *Manual of Procedures for Forest Mapping and Damage Detection using Satellite Data, Rep IRS-UP/SAC/FMDD/TN/16/90* Space Appl Center Ahmedabad (1990) 58

13 N V M Unni, M S R Murthy, S L Swamy, C B S Dutt *Fuelwood Sustainability in Rural Areas—A Model Study using Remote Sensing Data NRSA Tech Rep* (1992) 24

14 N V M Unni *Proc of Special Current Event Session* 41*st IAF Congress*, Dresden, Germany (1990) 49

15 Inderjit Singh *Photonirvachak* **16** (1988) 43

16 J S Parihar, P C Kotawal, S Panigrahi and N Chaturvedi *Study of Kanha National Park Tech Rep ISRO SP*-17-86 (1986) 65

17 P S Roy, R N Kaul and S S Gabrul *Int J Remote Sensing* **6** (1985) 411

18 P S Roy *Geocarto Int* **1** (1987) 19

19 C B S Dutt, M S R Murthy, S L Swamy, C S Jha and Kavitha Sharma *Identification of Grasslands in Western Ghats—A Pilot Study using Remote Sensing* NNRMS ISRO Hqs Bangalore (1993)

20 N V M Unni, P S Roy and V Parthasarathi *Int J Remote Sensing* **6** (1985) 419

21 P S Roy *Pilot Study on Grassland and Plantations* NRSA, Hyderabad (1993) (Unpublished)

22 C B S Dutt *Interface* **3** (1992) 7

23 D S Kamat and S K Susha *Proc Sem Crop Growth Conditions Remote Sensing* New Delhi (1984) 155

24 N Ravi, N K Patel, R R Navalgund, S Panigrahi and R N Dash *Photonirvachak* **20** (1992) 43

25 S Panigrahi, J S Parihar, N K Patel *Photonirvachak* **20** (1992) 35

26 S A Sharma, H P Bharr, Ajai and S Nanavathy *Photonirvachak* **19** (1991) 59

27 V K Dadhwal, D S Ruhal, T T Medhavy, S D Jarwal, A P Khera, J Singh, T Sharma and J S Parihar *Photonirvachak* **19** (1991) 1

28 D S Ruhal, T T Medhavy, S D Tarwar, V K Dadhwal, A P Khera and J S Parihar In: *Proc natn. Symp Remote Sensing for Agriculture Applications* New Delhi (1990) 250

29 Space Application Center (SAC) *Review Document of RSAM Project: Crops Acreage and Production Estimation* SAC Ahmedabad (1990) 125

30 T T Medhavy, Tara Sharma, R P Dubey, R H Hooda, K E Mothi Kumar, M Yadav, M C Manchanda, D S Ruhal, A P Khera and S D Jarwal *Photonirvachak* **21** (1993) 21

31 M B Potdar, R Sharma, R C Dubey and B C Biswas *Photonirvachak* **19** (1991) 45

32 U R Rao *Space Technology and Forest Management with Specific Relevance to Developing Nations in Proc of* 41*st IAF Congress Proc Sp Current Event Session* 41*st IAF Congress* Dresden, Germany (1990)

33 S Panigrahi and J S Parihar *Int J Remote Sensing* **13** (1992) 2943

34 M C Pereira and A W Setzer *Int J Remote Sensing* **14** (1993) 583

35 R Nelson and N Horning *Int J Remote Sensing* **14** (1993) 1463

36 J H Rouse, R H Hass, J A Schell and D W Deering *Third ERTS Symp Monitoring Vegetation System in Great Plains with ERTS* **1** (1973)

37 C R Perry Jr and L F Lautenschlager *Remote Sensing Environ* **14** (1984) 169

38 A J Richardson and C L Wiegand *Photogramme Engng Remote Sensing* **43** (1977) 1541

39 P Ashburn The vegetation index number and crop identification *Proc LACIE Symp* Johnson Space Center Houston USA (1978) 843

40 R D Jackson, P N Slater and P J Pinter Jr *Remote Sensing Environ* **13** (1983) 187

41 M Shibayama, C L Weigand and AJ Richardson *Int J Remote Sensing* 7 (1986) 233

42 R J Kauth and G S Thomas *The Tasselled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as seen by LANDSAT Proc Symp Machine Processing Remote Sensing Data* (1976) 41

43 L Hay, C A Kuretz, J B Odenweller, E J Scheffner and B Wood *Development of AI procedures for dealing with the effect of episodal events on crop temporal response*, Agristars Rep **SR-B9-00434** Johnson Space Center Houston USA (1979)

44 E C Crist and R C Cicone *Remote Sensing Environ* **14** (1984) 235

45 R C Cicone and M D Metzler *Remote Sensing Environ* **14** (1984) 257

46 A R Huete *Remote Sensing Environ* **25** (1988) 295

47 F Baret, G Guyot and D Major TSAVI: Vegetation index which minimises soil brightness effects on LAI and APAR estimation *Proc IGRASS*-89 **3** (1989) 1355

48 Li Y Demetriades, T H Shah, E T Kanemasu, J K Shultis and M B Kirkham *Remote Sensing Environ* **44** (1993) 81

49 F Baret, S Jacquemoud and J F Hanocq *Remote Sensing Rev* 7 (1993) 65

50 M Derrien, B Franki, M Legleau and A Sairounc *Int J Remote Sensing* **13** (1992) 1787

51 T Phulpin, J P Jullien and D Lasselin *Int J Remote Sensing* **10** (1989) 869

52 J R G Townshend and C O Justice *Int J Remote Sensing* 7 (1986) 1435

53 J U Hielkema, S D Price and W L Astle *Int J Remote Sensing* 7 (1986) 1499.

54 C O Justice, J R G Townshend, B N Holben and C J Tucker *Int J Remote Sensing* **6** (1985) 1271

55 C J Tucker, J A Gatlin and S R Schneider *Photogramm Engng Remote Sensing* **50** (1984) 53

56 D B Botkin, J E Estes, R M MacDonald and M V Wilson *Bioscience* **34** (1984) 508

57 D Lo, Seen Chong, E Mougin and J P Gastellu-Etchegorry *Int J Remote Sensing* **14** (1993) 1517

58 C J Tucker and P J Sellers *Int J Remote Sensing* 7 (1986) 1395

59 C J Tucker, C Vanpraet, E Boerwinkel and A Gaston *Remote Sensing Environ* **13** (1983) 461

60 D F McGinnis and J D Tarpley *Adv Space Res* **5** (1985) 359

61 C J Tucker, J R G Townshend and T E Goff *Science* **227** (1985) 369

62 S R Schneider, D F McGinnis and G Stephens *Int J Remote Sensing* **6** (1985) 59

63 L Hay and A P Cracknell *Int J Remote Sensing Soc Reading* UK (1984) 181

64 C J Tucker and P J Sellers *Int J Remote Sensing* 7 (1986) 1395

65 S T Chari *Satellite-based Drought Surveillance in India—Some Achievements. 12th Canadian Symp Remote Sensing* Vancover Canada (1989) 1961

66 R K Gupta *Adv Space Res* **12** (1992) 87

67 G H Rao, R K Gupta, T S V Nadham and S Prasad *Adv Space Res* **13** (1993) 249

68 IGBP *Improved Global Data for Land Applications: A Proposal for a New High Resolution Data Set* IGBP Rep No 20, Stockholm Sweden (1992) 87

69 R K Gupta, G H Rao, D Vijayan, T S V Nadham, Rajesh Sharma and S Prasad *Indian J Radio & Space Phys* (1993) (Accepted)

70 R K Gupta *Int J Remote Sensing* **14** (1993) 53

71 R K Gupta, S Prasad, T S V Nadham and G H Rao *Adv Space Res* **13** (1993) 261

72 R K Gupta *Int J Remote Sensing* **13** (1992) 515

73 R K Gupta *Int J Remote Sensing* **13** (1992) 715

74 R K Gupta *Adv Space Res* **13** (1993) 223

75 N Che and J C Price *Remote Sensing Environ* **41** (1992) 19

76 J L Roujean, M Leroy, A Podaire and P Y Deschamps *Int J Remote Sensing* **13** (1992) 685

77 A T Jayaseelan and S T Chari *Int J Remote Sensing* **14** (1993) 603

78 D W Deering and T F Eck *Int J Remote Sensing* **8** (1987) 89

79 M J Duggin and R W Saunders *Remote Sensing of A Cloudy Atmosphere* (*Ed* A Henderson-Sellers) Taylor Francis (1984) 241

80 F W T Penning, De Vries and H H Van Larr (*Eds*) *Simulation of Plant Growth and Crop Production*, Simulation Monographs, Pudoc, Wageningen The Netherlands (1982) 308

81 J G P W Clevers *Remote Sensing Environ* **29** (1989) 25

82 J G P W Clevers and W Verhoef *Remote Sensing Rev* 7 (1993) 43

83 J C Price *Remote Sensing Environ* **41** (1992) 29

84 J A Gamon, J Penuelas and C B Field *Remote Sensing Environ* **41** (1992) 35

85 A J Richardson, C L Wiegand, D F Wanjura, D Dusek and J L Steiner *Remote Sensing Environ* **41** (1992) 71

86 F Baret, G Guyot, A Begue, P Maurial and A Podaire *Remote Sensing Environ* **26** (1988) 213

87 C J Tucker *Remote Sensing Environ* **10** (1990) 23

88 S N Goward, D Dye, A Kerber and V Kalb *Geocarto Int* **2** (1987) 27

89 R R Nemani and S W Running *J appl Met* **28** (1989) 276

90 S N Goward and K F Heummrich *Remote Sensing Environ* **39** (1992) 119

91 S W Running and R R Nemani *Remote Sensing Environ* **24** (1988) 347

92 J D Tenhumen, J D Hesketh and D M Gates *Leaf Photosynthesis Models, in Predicting Photosynthesis for Ecosystem Models* (*Eds* J D Hesketh and J W Jones), C·R C Press Boca-Raton **II** (1980) 17
93 P J Sellers *Int J Remote Sensing* **6** (1985) 1335
94 PJ Sellers and J L Dorman *J appl Met* **26** (1987) 622
95 K P Gallo and C C Daughtry *Atmos J* **78** (1986) 752
96 S W Running *Remote Sensing Biosphere* **15** (1990) 1300
97 A S Hope, D E Petzold, S N Goward, R M Ragan *Water Resource Bull* **22** (1986) 1011
98 S W Running, R R Nemani, D L Peterson, L E Band, D F Potts, L L Pierce and M A Spanner *Ecology* **70** (1989) 1090
99 F L Barnet and D K Thompson *Remote Sensing Environ* **12** (1982; 509
100 H C Penman *Natural Evapotranspiration from Open Water, Bare Soil and Grass Proc R Soc Lond* **A193** (1948) 120
101 Mary S McKenney and Norman J Rosenberg *Agriç Forest Met* **64** (1993) 81
102 C W Thornthwaite *Geogr Rev* **38** (1948) 55
103 H F Blaney and W D Criddle *Determining Water Requirement in Irrigated Areas from Climatological and Irrigation Data* USDA Soil Conservation Service Technical Paper No. 96 (1950)
104 G H Hargreaves *Trans ASAE* **17** (1974) 701
105 Z A Samani and G H Hargreaves Int Irrig Center Utah State Univ Logan UT (1985) 105
106 M E Jensen and H R Haise *J Irrig Drainage Div ASCE* **89** (1963) 15
107 C H B Priestley and R J Taylor *Mon Wea Rev* **100** (1972; 81
108 J L Monteith *Evaporation and Environment, The State and Movement of Water in Living Organisms*, 19th Symp Soc Exp Biol Academic Press NY (1964) 205
109 M S McKenney and N J Rosenberg *Agric Forest Met* **64** (1993) 81-110
110 R G Allen, M E Jenson, J C Wright and R D Burman *Agron J* **81** (1989) 650
111 J L Monteith *J R Met Soc* **107** (1981) 749
112 R D Jackson *Adv Irrig* **1** (1982) 43
113 B Seguin In: *Meteorology and Environ Sciences* ICTP, Trieste (Italy) World Scientific Singapore (1990) 565
114 R D Jackson, R J Reginto and S B Idso *Water Resource Res* **13** (1977) 651
115 B Seguin and B Itier *Int J Remote Sensing* **4** (1983) 371
116 T J Schmugge, F Becker and Z L Li *Remote Sensing Environ* **35** (1991) 95
117 A A Van De Griend and M Owe *Int J Remote Sensing* **14** (1993) 1119
118 P J Sellers, F G Hall, D E Strebel, R D Kelly, S B Verma, B L Markham, B L Bald, D S Schimel, J R Wang and E Kanemasu *Rep FIFE: Experiment Operations and Interim Results* (1990) 228
119 A Aman, P H Podaere and R Frouin *IEEE Trans Geosci &Remote Sensing* **30** (1992) 326
120 N A Quarmby, M Milne, T L Hindle and N Silles *Int J Remote Sensing* **14** (1993) 199
121 J J Settle and N A Drake *Int J Remote Sensing* **14** (1993) 1159
122 R K Gupta *Potential of Satellite-Based Systems for IGBP Studies* Nat Symp IGBP, Madras India (1993) 10
123 J R G Townshend and C O Justice *Int J Remote Sensing* **9** (1988) 187
124 M Raffy *Remote Sensing Environ* **40** (1992) 101
125 D S Kimes, P R Harrison and P A Ratcliffe *Int J Remote Sensing* **12** (1991) 1987
126 R E Dickinson *Modeling Evapotranspiration for Three-Dimensional Global Climate Models in Climate Processes and Climate Sensitivity* (*Eds.* J E Hanson and T Takatrashi) Geophysical Morphology **29** (1984) 58
127 P J Sellers, Y Mintz, Y Sud and A Dalcher *J atmos Sci* **43** (1986) 505
128 *IGBP Report A Study of Global Change* **12** (1990) 232
129 *IGBP Rep Plant-Water Interaction in Large-scale Hydrological Modeling* No 17 (1991) 44
130 NASA *Earth Observations Systems (EOS) Reference Hand Book* (*Eds* G Asrar and D J Dokken) National Aeronautics and Space Administration Washington D C USA (1993) 145
131 V V Solomonson and D L Toll *Adv Space Res* **11** (1991) 231
132 NASA, *Earth Observations Systems (EOS) Reference Hand Book* (*Ed.* D J Dokken) National Aeronautics and Space Administration Washington DC USA (1991) 147

SATELLITES FOR MONITORING CLIMATE CHANGE: THE EMERGING SCENARIO

R R KELKAR
India Meteorological Department, New Delhi-110 003, *India*

(*Received* 12 *July* 1993; *Accepted* 1 *September* 1993)

Although in the last few years, large-scale utilisation of satellite data has become possible for climate studies, there is also an increasing realisation of the limitations of the processes by which information from meteorological satellites is being converted to a climate-scale data base. This has led meteorologists and satellite planners to think very seriously about using modified or totally new wavelength channels, increasing the number of channels of radiometers, improving the resolutions, trying new orbits, etc. Some of these plans are very ambitious and if brought into reality, they are likely to completely alter the satellite scenario by the turn of the century. This paper first reviews the current status of various on-going programmes and efforts and then describes the new possibilities that are opening up, based upon the plan projections of various countries available presently.

Key Words: Satellites; Monitoring Climate Change; Wavelength Channel; Temperature Retrievals; Earth Radiation Budget; Ozone Hole; Vegetation; Snow Cover; Microwave Remote Sensing

Introduction

The primary objective of the operational weather satellite programmes of various countries has been to augment the conventional inputs available to weather forecasters. Satellites like INSAT, GOES, METEOSAT or TIROS have all carried high-resolution radiometers designed to measure infrared radiances from cloud tops or the earth's surface in spectral windows which have minimal absorption. The main products from these measurements are cloud imagery and sea surface temperature/cloud top temperature information, with cloud motion vectors derived as a useful by-product. METEOSAT was the first satellite to provide water vapour channel imagery, while the U.S. polar-orbiting satellites introduced the capability of radiance measurements in 20 different wavelength channels for sounding purposes.

Over the years, and in the last decade in particular, many efforts have been launched to make use of meteorological satellite data in a manner aiming far beyond the immediate requirements of weather analysis and forecasting. Large-scale precipitation is being derived from infrared channel temperature histograms under the Global Precipitation Climatology Project (GPCP), while cloud climatologies on different scales are being generated under the International Satellite Cloud Climatology Project (ISCCP). Outgoing Longwave Radiation (OLR) derivations from meteorological satellites are finding an important place in earth radiation budget studies.

The derived products like OLR, precipitation, moisture, etc. are being used in validating atmosphere-ocean models, in climate simulation experiments and for diagnostic studies of the climate system. Although the record lengths of these products are much shorter (10-20 years) than conventional climatological data, and there are problems like sensor degradation, calibration errors, etc. the extensive coverage in space and time provided by satellite products enhances their usefulness.

Although in the last few years, large-scale utilization of satellite data has become possible for climate studies, there is also an increasing realisation of the limitations of the processes by which information from meteorological satellites is being converted to a climate-scale data base. This has led meteorologists and satellite planners to think very seriously about using modified or totally new wavelength channels, increasing the number of channels of radiometers, improving the resolutions, trying new orbits, etc. Some of these plans are very ambitious and if brought into reality, they are likely to completely alter the satellite scenario by the turn of the century. This paper first reviews the current status of various on-going programmes and efforts and then describes the new possibilities that are opening up, based upon the plan projections of various countries available presently.

Climate Monitoring through Weather Satellites

This section gives a description of past and current efforts to construct long-term data sets of parameters crucial to climate monitoring, using data available from meteorological satellites.

Precipitation Measurements from Satellites

Under the World Climate Research Programme (WCRP), the World Meteorological Organization (WMO) has established the Global Precipitation Climatology Project (GPCP) whose objective is to prepare a climatology of precipitation over the globe in order to achieve a better understanding of the hydrological cycle and energy budget of the atmosphere. The importance of the GPCP programme lies in the fact that a large fraction of the energy input from the earth's surface to the atmosphere is in the form of latent heat of water vapour and its subsequent release by condensation in the atmosphere. Moreover, the global patterns of precipitation constitute the most significant manifestation of climate and are the most sensitive indicators of climate change.

The GPCP Project started in 1986 and is to run upto 1995. It plans to use both geostationary and polar-orbiting satellite data during this 10-year period in conjunction with surface data in order to prepare the precipitation climatology. The satellite precipitation retrieval method uses 16-class temperature histograms, prepared daily and accumulated over 5-day periods[1].

Computations of precipitation on a 2.5° × 2.5° scale over the Indian and Indian Ocean region have been made with INSAT data regularly over the last few years. Such estimates have proved to be useful in studying the interannual variability of monsoon rainfall[2].

Ferraro[3] has indicated that for the purposes of climate studies, precipitation measurements from satellites are required on a 50 × 50km scale with an accuracy of 1mm/day or 10%, whichever is smaller. However, monthly totals with spatial resolutions of 2.5° × 2.5° are also needed for studies of interannual variability. Researchers are also seeking an improvement to 25×25km, 6 hourly and 10% accuracy as desirable goals.

In the GPCP-type algorithms, estimates of tropical rainfall based upon the IR temperature threshold technique are useful, but significant biases have been found particularly those associated with non-raining cirrus and rain from warm clouds[4]. To correct for these biases, locally applicable calibrations would have to be developed. In general, microwave channel algorithms have produced instantaneous rainfall estimates which are superior to those of the IR-based algorithms.

Hybrid alogirthms using both infrared and microwave brightness temperature have been found to have smaller sampling errors and improved estimates of precipitation for both weather and climate applications[5].

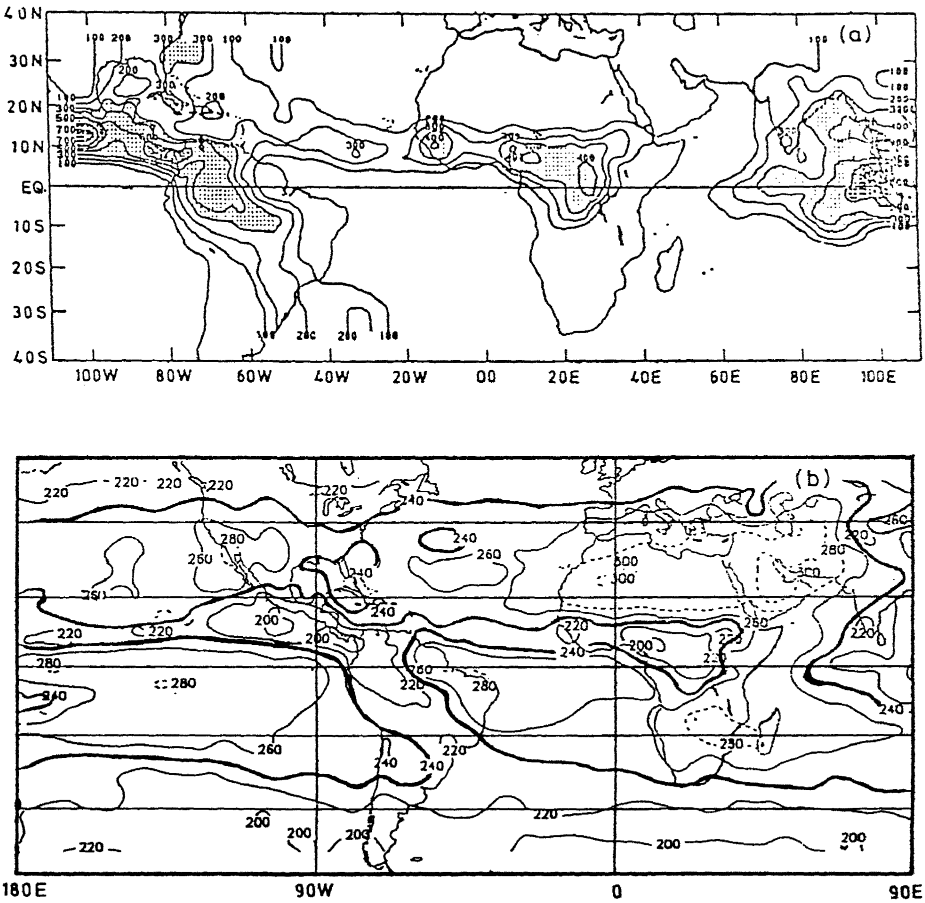


Fig 1 Monthly mean (*a*) Precipitation and (*b*) OLR derived from NOAA-11 AVHRR for September 1992. (*Source*: *Climate Diagnostic Bulletin*, NOAA/CAC).

Earth Radiation Budget

Since the early years of the weather satellite programme, there has been sufficient interest in the measurement of the earth's radiation budget components[6]. However, the quality of the derived data has been compromised by changes in satellites, orbits and sensors as well as the algorithms for correlating narrow band radiances to broad band values. In India also, derivations of Outgoing Longwave Radiation (OLR) are being routinely made from INSAT thermal infra-red data since 1986, on a 2.5°×2.5° grid over daily, weekly and monthly time scales. INSAT OLR is of potential value for studies of seasonal and inter-annual variations over the Indian Ocean region[2] as it is well-correlated with convective and precipitating regimes. Fig. 1 shows that there is good correspondence between OLR and large-scale precipitation on a monthly scale. It is seen that the $240 w/m^2$ OLR isopleth generally separates the raining areas from the non-raining ones.

The Earth Radiation Budget Experiment (ERBE) of the U.S. is a coordinated effort to use data from identical instruments on-board different satellites to measure the broad-band components of the earth's radiation budget to a high degree of accuracy with diurnal resolution[7]. The ERBE project involves three satellites, NOAA-9, NOAA-10 and ERBS (Earth Radiation Budget Satellite). ERBS is in a 57° inclined orbit at 600km altitude, which precesses through all hours of the day at the equator once every 36 days, thus providing a complete diurnal sampling. The ERBE scanning radiometers provide broadband measurements in both shortwave (0.2-5.0 μ) and longwave (5-50 μ) regions at 30km resolution. These are distinctly superior to the indirect OLR derivations from the 10.5-12.5 μ window measurements. The ERBE data collection began in 1985 and continued till Feb. 1990.

The NIMBUS-7 research satellite launched in 1978, in a local noon/local midnight orbit, has produced several long-term data sets. NIMBUS-7 had a Temperature Humidity Infrared Radiometer (THIR) and Wide Field of View Radiometer (WFOVR) measuring shortwave (0.2-3.8 μ) and total longwave (0.2-50 μ) radiation. NIMBUS-7 ERB data has coarse resolution, lower calibration accuracy, and a systematic degradation of the sensor with time. Its advantage lies in the consistent viewpoint and the longer period of data availability. Hurrel and Campbell[8] have shown various comparisons of OLR obtained from ERBE, NOAA and NIMBUS-7. One such comparison of the three OLR patterns for the month of July 1985 is reproduced in Fig. 2. The NIMBUS-7 pattern is seen to be smoother because of the large field of view of the sensor. Although the general spatial distributions agree, there are noticeable differences in the magnitudes as may be expected.

Ozone Hole

Ozone measurements through satellites are of three categories: (1) total ozone measurements by TOMS (Total Ozone Mapping Spectrometer) on the NIMBUS-7 satellite, (2) total ozone retrievals from TOVS (TIROS Operational Vertical Sounder) on NOAA satellites, and (3) Ozone profile measurements in 25-55km altitudes by SBUV (Solar Backscatter Ultraviolet Technique) instrument on NOAA and NIMBUS Satellites. The three data sets are available

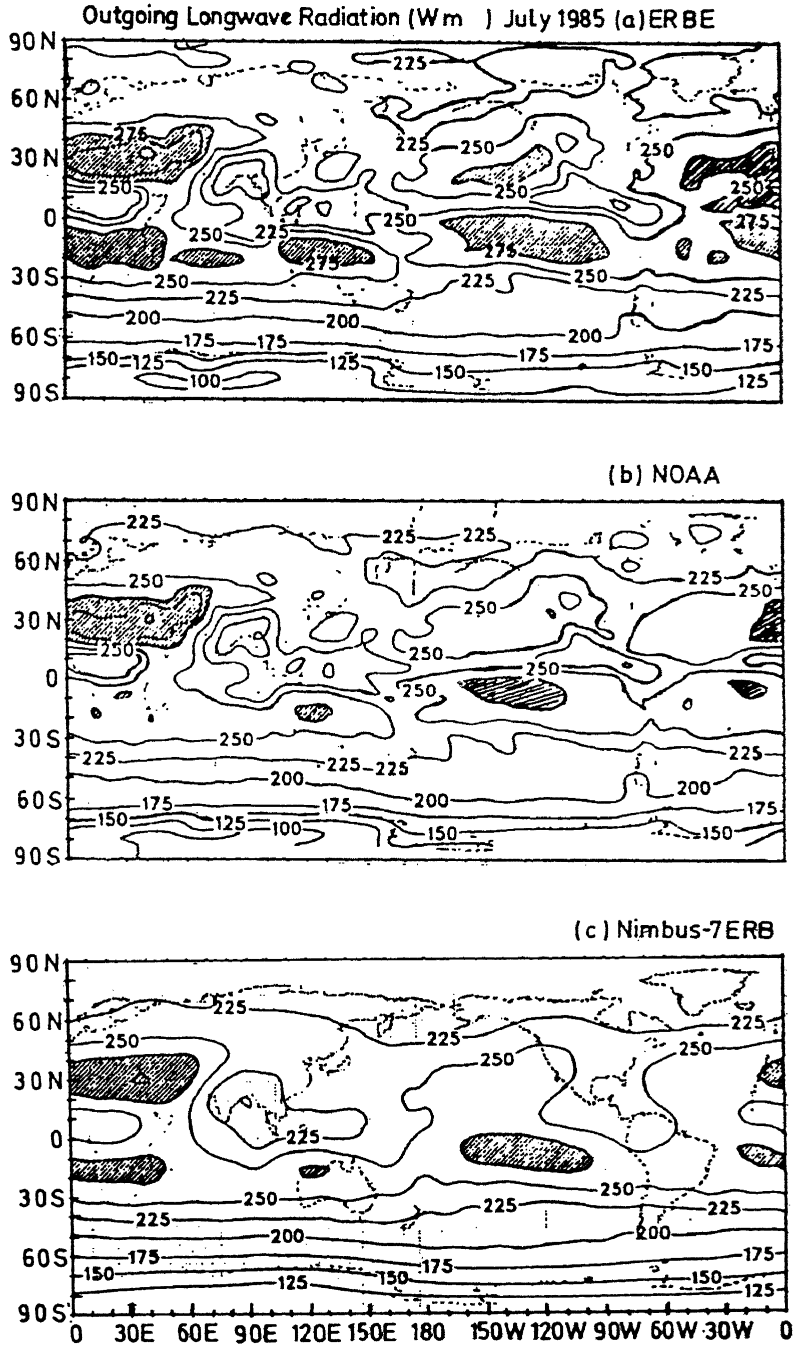


Fig 2 Outgoing longwave radiation in W m^{-2} for July 1985 from (*a*) ERBE, (*b*) NOAA and (*c*) Nimbus-7 ERB. Values less than 225 W m^{-2} are indicated by stippling and values greater than 275 W m^{-2} are hatched. The contour increment is 25 W m^{-2}.

from 1978, 1979 and 1985 respectively. The TOMS instrument uses differential absorption of reflected solar ultra-violet radiation to estimte total ozone, while TOVS brightness temperature in the 9.7 μ band is used for total ozone retrievals. The SBUV operates on the principle of detecting the spectral distribution of backscattered solar ultraviolet radiation. The horizontal resolution of TOVS is 17km, that of SBUV is 200km and TOMS is 100km. Each of these three ozone data sets thus have their own strengths and weaknesses.

Recent observations of ozone have indicated that there is a significant depletion of ozone in the vertical column during the spring season of the Southern Hemisphere (Sept-Oct) over the Antarctic region. This phenomenon, which has come to be known as the "ozone hole", has become a subject of intense study and investigation. Both the depth and area of the hole are now being monitored regularly through measurements made by TOVS. Fig. 3 taken from Halpert *et al.*[9] shows the time series of the area over which TOVS derived total ozone had a value of less than 212 Dobson Units (DU). Total ozone values of such low magnitude are not generally observed outside the southern polar regions. Even over the polar regions, they were never so low before the eighties. The threshold of 212 DU therefore is a convenient mark for delineating the ozone hole boundary. From late September to early October 1992, the ozone hole was seen to have its greatest areal extent so far, perhaps due to the presence of aerosols from the Mt. Pinatubo volcanic eruption of June 1991. A further gradual depletion of ozone is expected to occur in years to come as the concentration of stratospheric chlorine and bromine would increase.

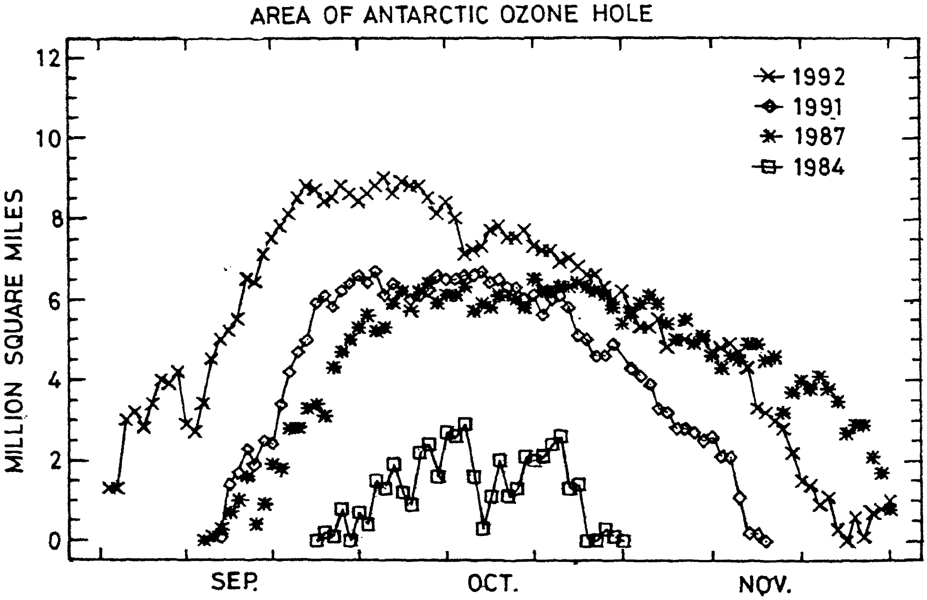


Fig 3 Time series of the area (million square miles) over which TOVS total ozone was measured to be less than 212 Dobson Units for the September-December period for 1984, 1987, 1991 and 1992

Aerosols

The volcanic eruptions of El Chichon in March 1982 and Mt. Pinatubo in June 1991 have had effects which could be noticed in satellite imagery and satellite-derived temperature products. In such events, sulphur dioxide gets injected into the stratosphere where it slowly combines with water vapour to form sulphuric acid. This in turn leads to more solar radiation being reflected to space. The Aerosol Optical Thickness (AOT), which controls the reflection of solar radiation back to space, is typically less than 0.1, except in areas having dust, smoke, haze, industrial waste products, etc. However, following Mt. Pinatubo's eruption, AOT has been found to exceed the value of 0.1 over a large part of the earth's surface. An increase in AOT by 0.1 corresponds to a 1% increase in the earth's albedo. This is an indirect measure of the cooling effect on the earth's climate by factors such as stratospheric loading by volcanic particles. Since 1987, NOAA/AVHRR Channel 1 radiance measurements are being used to derive AOT and time series of departures of AOT from mean values provide an insight into the transport of aerosols within the earth's atmosphere and its impact on climate[10]

Vegetation and Snow Cover

In 1982, NOAA began producing weekly global vegetation index maps primarily as a data source for assessment of global agricultural production. Normalised Difference Vegetation Index (NDVI) is also being computed on a finer local scale over the Indian Region by the National Remote Sensing Agency[11]. Since recently, vegetation indices are finding use in climate monitoring, and also as boundary conditions in climate prediction models. However, the usefulness of an NDVI climatology is uncertain because of biases introduced by cloud contamination and scan angle differences in the short term and orbital drift and sensor degradation in the long term. These problems need to be eliminated before interannual variability of NDVI can be gainfully used in climate studies.

Since 1966, NOAA has been bringing out Northern Hemisphere Weekly Snow and Ice Cover Charts based upon manual analysis of satellite visible imagery. This effort has resulted in the compilation of a long-term data set[12] which is useful for monitoring the continental snow cover and for deriving snow cover anomalies. Satellite-derived fields of snow depth are also available in areas which are not heavily vegetated. According to current thinking, cryospheric variations of the higher latitudes will have a key role to play in future climate and global change and such parameters therefore need to be incorporated with precision in climate prediction models. The Eurasian snow cover is also an important parameter used operationally in the long-range forecasting of the Southwest monsoon rainfall[13].

Temperature Retrievals

The three channels of the Microwave Sounding Unit (MSU) of NOAA satellites measure a vertically averaged thermal emission by molecular oxygen in the 60GHz band. Oxygen is uniformly mixed in the atmosphere and its concentration is very stable in time. The MSU temperatures are, therefore, suitable for monitoring intra-seasonal and inter-annual temperature variations of the troposphere and stratosphere[14].

The MSU Channel 2 peaks near 500hpa and is dominated by the tropospheric temperature signal, while channel 4 which peaks at 70hpa is representative of stratospheric temperatures. The MSU brightness temperatures have been carefully calibrated over more than 13 years and because of their interconsistency, it is possible to compute anomaly fields. The anomalies have been found to reflect the effects of volcanic eruption in the lower straqtosphere and ENSO events on the tropospheric temperatures.

Although the retrieval of sea surface temperature from the thermal window radiances measured by satellites is a simple process, the accuracy of the retrieval is utilised by the effects of the atmospheric absorption which cannot be parameterised and therefore not accounted for. The best way is to eliminate them from the considerations of the retrieval by using multi-channel measurements. The adoption of the Multi-Channel Sea Surface Temperature (MCSST) algorithm by NOAA in 1981 resulted in greatly improved SST retrievals, comparing within 0.4°C of drifting buoy matchups. However, the algorithms need to be constantly fine-tuned and there are other problems like paucity of *in situ* data over southern hemispheric oceans and the inability of satellite techniques to retrieve SST in cloudy areas, which is a big constraint over the monsoon regime.

Future Possibilities

Microwave Remote Sensing

After over 25 years of meteorological remote sensing in the solar and infrared channels, there has been a sudden upsurge of interest in the exploitation of the microwave frequencies for meteorological purposes. The Defence Meteorological Satellite Programme of the U.S. (DMSP) launched in 1987, a polar-orbiting low altitude satellite having the Special Sensor Microwave Imager (SSM/I). This very successful mission has been followed up by more microwave instruments in the DMSP programme. Moreover, the utility of the data has been greatly enhanced because the sharing of the data with civilian users in the U.S. was permitted, thus opening up a new field of applications[15].

The SSM/I is a passive system measuring microwave radiation in the 19, 22, 37 and 85GHz bands with resolutions ranging from 69×43 to 15×13km. The SSM/T instrument (Special Sensor Microwave/Temperature) is a cross-track scanning radiometer with seven channels in the 50-60GHz region peaking at various heights from the surface to 30km. For water vapour soundings, a different version of SSM/T, called SSM/T2 will fly on future DMSP Satellites. This will be followed by SSM/IS (Special Sensor Microwave/Imager Sounder), a conical scanner, incorporating the channels of all the three instruments, SSM/T, SSM/T2 and SSM/I.

The NOAA K-N satellites will carry an Advanced Microwave Sounding Unit, AMSU/A, which will be a 15-channel radiometer capable of providing temperature soundings from the surface to 3hPa with a 50km nadir resolution.

AMSU/B will be a 5-channel radiometer with 15km resolution, for derivation of water vapour profiles.

The data from the SSM's and AMSU's will help in the estimation of global precipitation over land and water, global cloud liquid water content over water, and water vapour and atmospheric temperature profiles. Studies indicate that for climate monitoring, the best results are obtained when data from different channels like visible, infrared and microwave are used in conjunction, so that the shortcomings of one are largely overcome by the advantages offered by the others. This is particularly so in the case of global precipitation, where infrared-plus-microwave algorithms have yielded promising results. Here the twice a day microwave data are matched with 8 times a day infrared data from geostationary satellites, while the physically-based microwave rainfall estimates are used to calibrate or fine-tune the indirect infrared estimates.

Tropical Rainfall Measuring Mission (TRMM)

The Tropical Rainfall Measuring Mission (TRMM) is a joint project of the U.S. and Japan for making extensive and accurate measurements of tropical rainfall[16]. The TRMM satellite, which will be launched in 1997 is envisaged to be in a 35° inclination orbit. It will be non-sun-synchronous, so as to provide sampling at different times of the day with a repeat cycle of about 3 days. It will have a low altitude of about 350km.

TRMM will combine the advantages of (*a*) an AVHRR type instrument viewing in the solar and thermal infrared channels for indirect precipitation estimation, (*b*) a microwave imager of the SSM/I type operating at 19, 22, 37 and 85 GHz to provide physical rainfall retrievals, (*c*) a scanning microwave radiometer at 19GHz, and (*d*) a precipitation radar working at a frequency of 14GHz with a footprint of 5km. This combination of data from both active and passive microwave instruments with simultaneously available visible and infrared measurements is expected to lead to a quantum jump in precipitation monitoring over the tropics.

Geostationary Satellites

A major design change in the next-generation U.S. geostationary weather satellites of the GOES I-M Series will be that they will be three-axis stabilised spacecraft similar to INSAT[17]. With a spinning satellite in geostationary orbit, the sensor sees the earth for less than 5% of the time. In a three-axis stabilised configuration, it can be designed to view the earth for upto 75% of the time, thus improving greatly the radiometric efficiency.

GOES I-M will have a 5-channel imager (1 visible and 4 *i-r*) scanning the earth's disc in 30 minutes[18]. The resolution at the subsatellite point will be 1km for visible, 4km for *i-r* and 8km for the water vapour channels (Table I) They will also have a 19-channel sounder (1 visible and 18 *i-r*) comparable in performance to the present NOAA TOVS with resolution of 8km and a repeat time of 42 mintues (Table II).

In the GOES-N system, which will follows GOES I-M, there is a proposal to have two imaging instruments working parallely, one seeing the full disc and another for faster scanning of small sectors. GOES-N may also have 5 new channels at 4km resolution, and the resolution of the GOES I-M channels may be improved by a factor of two. The sounder accuracy and resolution are also planned to be improved.

Table I

GOES I-M Imager channels

Channel Number	Wavelength Range (μ)	Band	Resolution (km)	Application
1	0.55-0.75	Vis	1	Daytime cloud cover
2	3.80-4.0	*i-r*	4	Night-time cloud cover
3	6.50-7.0	*i-r*	8	Water vapour imaging
4	10.20-11.20	*i-r*	4	Surface temperature
5	11.50-12.50	*i-r*	4	SST, water vapour

Table II

GOES I-M sounder channels

Channel Number	Central Wavelength (μ)	Resolution (km)	Application
1	14.71		Temperature sounding
2	14.37	8km for all channels	,,
3	14.06		,,
4	13.96		,,
5	13.37		,,
6	12.66		,,
7	12.02		Surface temperature
8	11.03		,,
9	9.71		Total ozone
10	7.43		Moisture sounding
11	7.02		,,
12	6.51		,,
13	4.57		Temperature sounding
14	4.52		,,
15	4.45		,,
16	4.13		,,
17	3.98		Surface temperature
18	3.74		,,
19	0.70		Cloud imaging

Table III

MSG SEVIRI channels

Band	Wavelength Range (μ)	Resolution (km)	Application
High-resolution visible	0.50-0.90		High-resolution imaging
Visible	0.56-0.71		Basic imaging
,,	0.71-0.95		
Infrared	1.44-1.79		
,,	3.40-4.20		
,,	8.30-9.10		
,,	9.80-11.80		
,,	11.00-13.00		
Water vapour	5.35-7.15		Air mass monitoring
,,	6.85-7.85		
Ozone	9.46-9.94		
Carbon dioxide	13.04-13.76		

The current European meteorological satellites were initially designed in the early seventies and have been used in orbit since 1977. These are planned to be replaced by the Meteosat Second Generation (MSG) satellites by the year 2000. The MSG mission, besides providing continuity of present services, envisages support to climate and environmental monitoring as well as to earth resources management. The basic instrument to be carried by MSG is the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI), consisting of a 7-channel imaging radiometer, a 4-channel air mass monitor and a single-channel high resolution visible imager (Table III). The air-mass monitor will have two channels in the water vapour absorption band and one each in the carbon dioxide and ozone bands. All these channels of SEVIRI will have a resolution of 3km at the sub-satellite print and a scanning cycle of 15 minutes for the full disc, except the high resolution visible channel which will have a better resolution of 1km at the sub-satellite point with a 15 min. scan time[19].

Japan proposes to launch its geostationary meteorological satellite GMS-5 in 1994 which will be same as earlier GMS satellites excepting that a water vapour channel is proposed to be added. The Indian geostationary satellite, INSAT which is a multi-user system, will not have meteorological payloads in the INSAT-2C and INSAT-2D spacecrafts. The meteorological component will be restored in INSAT-2E which will be the last satellite of the INSAT-2 series.

U.S./European Polar Orbiting Satellites

As a replacement of the currently operational NOAA-11 and NOAA-12 polar orbiting satellites, the U.S. plans to launch what has been termed as the NOAA K-N series[18]. In these satellites, the AVHRR will be upgraded with the addition of a sixth channel at 1.6 μ to provide better discrimination of snow and ice from clouds. The MSU and SSU will be replaced by the Advanced Microwave Sounding Unit (AMSU) which would have 20 microwave channels at 45km nadir resolution.

A further significant departure from what has been the operational policy of the U.S. for a long time, the U.S. proposes to have only one (in place of two) polar orbiting satellite at a time, starting from NOAA-O. This will provide coverage at local afternoon (1330) in the ascending node. For the coverage at local morning (0730) in the descending node, the U.S. has sought the partnership of Eumetsat to provide an identical satellite, which has been termed as the Eumetsat Polar System[20]. The complementary system of NOAA and EPS is expected to become operational by the year 2001. Both the satellites are likely to have a 7-channel, 1.1km-resolution imager and more advanced infrared and microwave sounders, in addition to a host of other instruments like SBUV, TOMS, altimeter and scattermeter, etc.

The sophistication planned to be introduced in these new polar-orbiting satellites will lead to the availability of measurements of a wide range of atmospheric as well as oceanographic variables such as precipitation rates over ocean and land, cloud water and vapour content, snow/ice cover, radiation budget components and all-weather retrievals of temperature profiles, SST, ocean surface winds, wave heights, vegetation, trace gas concentrations, etc.

Continued Measurements of Total Ozone

The Total Ozone Mapping Spectrometer (TOMS) on-board NIMBUS-7 has provided total column ozone over the globe every day since 1978. Later in this decade, the TOMS instrument will fly on other satellite systems, first on the Russian METEOR-3, then on Japan's ADEOS and after that on the Earth Probes Programme of the U.S.

This cooperation and coordinated activity of three different countries will ensure continued availability of total ozone data to meet the needs of climate change monitoring beyond the life of NIMBUS-7.

The Upper Atmospheric Research Satellite (UARS) launched by the U.S. is also providing data on variations in global ozone and the dynamics of the upper atmosphere.

Other Satellites and Instruments

The European Remote Sensing Satellite ERS-1, launched in 1992, and the other satellites of this type which will be launched by Japan (JERS), are primarily earth resources satellites. However, the ERS carries the Synthetic Aperture Radar (SAR) which is capable of measuring ocean surface topography and deriving ocean surface winds. The repeat cycle of ERS-1 is 3 days for ocean surface wind measurements. Such measurements are not possible from geostationary satellites by the cloud displacement technique below 900hPa level.

Japan[21] is also planning the Advanced Earth Observation Satellite (ADEOS), which will have an Ocean Column and Temperature Scanner (OCTS) with 6 visible channels and 6 *i-r* channels at a surface resolution of 0.7km. It will also have an Advanced Visible and Near Infra-red Radiometer

(AVNIR) with 3 visible channels, 1 near *i-r* channel and a panchromatic channel. Besides these, ADEOS is expected to have several specialised payloads. Among these will be a NASA Scatterometer (NSCAT) for measurement of ocean surface wind speed (±2m/s) and direction (±20°) over the globe. It will also have an Interferometric Monitor for Greenhouse Gases (IMG) for measurements of the concentration of CO_2, N_2O, CH_4, etc.) and a Retroreflector in Space (RIS) for measurement of Ozone, CFC's, CO_2, etc. by laser beam absorption technique, transmitted by a ground station and reflected back by ADEOS.

Mission to Planet Earth

The U.S. Mission to Planet Earth is a programme which is based on the realisation that the earth faces the prospects of accelerated environmental change including climate warming, rising sea levels, deforestation, desertification, ozone depletion, acid rain and reduction in biodiversity. It is also felt that present means for monitoring these changes are inadequate. The U.S. therefore proposes to have a comprehensive programme of global and climate change monitoring based upon sophisticated, Earth Observing System (EOS) satellites to be launched by U.S. in the years 2000 and beyond and supported by international satellite programmes during this decade itself. The existing programmes include UARS, TOPEX/Poseidon, NSCAT, ERBE, TOMS, etc. whereas the future missions planned are the NASA Earth Probes and Geostationary Platforms. The latter will include various payloads to monitor the hydrological, biogeochemical, climatic and geophysical processes within the total earth system. Much of the Mission to Planet Earth is currently in the stage of conceptual evolution[22] and the precise shape that it will take in the coming years remains to be seen.

Turn-of-the-Century Scenario

Although the availability of financial resources is the major limiting factor in the growth of space programmes, it is certain from the current enthusiasm that the environmental satellite scenario in the years 2000 and beyond will be totally different from what it was in the 1980's or 1990's. Besides satellites in geostationary and polar orbits, there would be satellites orbiting over the tropical belt and satellites in non-sun-synchronous orbits to provide total diurnal sampling. Cooperative arrangements amongst nations will result in identical instruments being flown by different countries (e.g., TOMS). The polar orbiting meteorological satellites will no longer be the sole responsibility of the U.S. but Europe will share half the burden. Microwave imagers and sounders will perhaps outnumber the conventional visible/infrared radiometers which themselves will undergo a distinct improvement. Retrieval algorithms will almost certainly make use of data from both infrared and microwave channels to improve the accuracy and temporal sampling. New parameters like concentration of greenhouse gases and CFC's will be monitored, while continuity will be provided for the present measurements of ozone, atmospheric temperatures, radiation balance components, etc. Measurements of sea surface temperature, surface vegetation cover, snow cover, etc. may no longer be restricted to non-cloudy conditions.

By all means, the turn-of-the-century satellite scenario holds out of prospect that will excite anyone. And it is also reassuring to know that the earth and space science community is geared to meet the challenges of global and climate changes, the consequences of which are sometimes made out to be so awesome.

References

1 P A Arkin and B N Meisner *Mon Wea Rev* **115** (1987) 51-74

2 R R Kelkar *Proc Indian Acad Sci (EPS)* **102** (1973)

3 R Ferraro (Ed.) *NOAA Tech Rep NESDIS* **62** (1992) 61pp

4 P A Arkin, A V R K Rao and R R Kelkar *J Climate* **2** 619-628

5 R A Adler, A J Negri and I M Hakkarinen *Global Planetary Change* **4** (1991) 87-92

6 A Gruber and J S Winston *Bull Am Met Soc* **59** (1978) 1570-1573

7 E F Harrison and 7 others *Bull Am Met Soc* **69** (1988) 1144-1151

8 J W Hurrell and G C Compbell *Tech Note NCAR/TN*-371-*STR* NCAR Boulder Colorado (1992) 94pp

9 M S Halpert and C F Ropelewski Fourth Annual Climate Assessment 1992 Climate Analysis Center NOAA (1993) 90pp

10 C R N Rao, L L Stowe and E P McClain Remote Sensing of Aerosols over the Oceans using AVHRR Data: Theory Practice and Applications *Int J Remote Sensing* **14** (1989) 743-749

11 R K Gupta *Int J Remote Sensing* **13** (1992) 715-735

12 M Matson, C F Ropelewski and M S Varnadore *An Atlas of Satellite-derived Northern Hemispheric Snow Cover Frequency U S Dept Commerce* (1986) 75pp

13 V Gowariker, V Thapliyal, R P Sarker, G S Mandal and D R Sikka *Mausam* **40** (1989) 115-122

14 R W Spencer and J R Christy *Science* **247** (1990) 1558-1562

15 R G Isaacs, E Kalnay, G Ohring and R McClatchey *Bull Am Met Soc* 74 (1993) 87-91

16 J Simpson, R F Adler and G R North A Proposed Tropical Rainfall Measuring Mission (TRMM) Satellite *Bull Am Met Soc* **59** (1978) 1570-1573

17 A Schwalb Envirosat-2000 Report: *GOES-NEXT Overview NOAA* Washington D C 38pp

18 H J Hussey, S R Schneider, R S Gird and B H Needham *Global Planetary Change* **4** (1991) 289-295

19 M Perrone Future SEVIRI Observations from Meteosat Second Generation Proc Ninth Meteosat Users Meeting EUM-P-11, Eumetsat Darmstadt Germany 379-387

20 J Morgan European Systems of Operational Mete Satellites *Global Planetary Change* **4** (1991) 297-301

21 K Maruo *ESCAP Remote Sensing Newslett* **10** (1993) 6-7

APPLICATION OF SATELLITE DATA FOR MONSOON RESEARCH

A V R KRISHNA RAO

Satellite Meteorology Division, India Meteorological Department,
Lodi Road, New Delhi 110 003, *India*

(*Received* 4 *August* 1993; *Accepted* 3 *September* 1993)

In this paper, the intraseasonal and diurnal variations of the cloudiness during monsoon season of 1991 are presented using a composite data set of INSAT & GMS satellites. The parameter chosen, to study these aspects, is the cloud top temperature (CTT) derived over every 1 deg lat./Long. grid, covering the area 60N-60S and 40E-160W. The composite data set was prepared after deriving a suitable regression relationship between the CTTs obtained through both the satellites, over the common area of view, equidistant from both of them.

The maps of horizontal distribution of the CTTs during the months May through September reflect the position of ITCZ very well. Since the mean CTTs indicate the mean cloud height particularly that of convective clouds in each 1 deg square mesh, the variations of CTTs reflect those of convection. The three hourly CTT distribution indicated an east-west fluctuation of convective clouds over the North Bay of Bengal and adjoining India with time. Lowest CTTs are found over north Bay at about noon, and they move westwards with time into the land upto 21Z and swing back to North Bay of Bengal by morning.

The diurnal variation of convection over different regions, also, was presented using 3 hourly mean CTT values.

A band-pass filter of 30-40 days was applied to the CTT anomalies, after removing the seasonal trend from the data set. These anomalies representing maximum cloud zone (MCZ) show, on time-latitude cross-section, a northward movement starting from the equator to about 30N over the Indian longitudes only, with an average phase-speed of 1 deg. lat./day. The average periodicity is found to be 40 days. Elsewhere the mode is not prominent and travels upto 20 deg. N only. The time-longitude cross-section of CTT anomalies show an eastward movement of the anomalies in the equatorial region. Each epoch of the northward movement of the MCZ coincides with a corresponding eastward movement of cloudiness in the equatorial region.

The time-longitude cross-sections of 3-hourly CTTs in different latitudinal belts indicate a westward movement of the clouds north on 5N, being maximum in July and in the latitudinal belt of 10-15N. These synoptic scale transient systems with associated convective clouds originate near about 160E longitude and move westwards. After entering the Bay of Bengal, most of them weaken west of 80E longitude. They travel with a phase speed of 5 deg. long./day and the average periodicity of their occurrence is found to be 9 days.

Key Words: Satellite Data; Monsoon Research; Cloud Top Temperature; Diurnal Variations; Maximum Cloud Zone; Cloud Motion Vector; Quantitative Precipitation

Introduction

The Indian Southwest Monsoon is a fascinating phenomenon. It means different to different people. The common man finds a great relief from the sweltering heat of the summer season with the onset of rains. While the farmer eagerly awaits the rains for his agricultural operations, the planner of the Country's economy closely monitor its performance since the economy of the country depends on good monsoon. For a meteorologist, engaged in the prediction of its phases, every monsoon throws up new problems as no two monsoons are similar in their performance. Although the monsoon circulation is described, as akin to that of Land and Sea breeze on a broader scale, the interactions between different scales of motion, within its circulation pattern, makes the problem more complex. Mesoscale circulations are strongly coupled with synoptic scale & planetary scale circulations through complex interactions which are not yet completely understood. Inspite of the unfailing regularity with which the Monsoon occurs year after year, the aberrations it exhibits in its various facets are still a challenging problem to meteorologists. Owing to the influence of the monsoon circulation on the Global circulation, the International Meteorological community conducted four International Experiments. (IIOE of 1963, ISMEX 73 & 77, and Monex 79). Though a wealth of quantitative information was collected through these experiments, the understanding of the monsoon system in its totality remained still elusive. These experiments, at least, helped in identifying the features which need to be monitored on a year to year basis. Extensive research work has been conducted on various facets of monsoon such as its onset, active and weak phases, break in monsoon, monsoon depressions and numerical studies. Excellent treatises on the research work conducted were presented in the Monographs by Rao[1] and Das[2].

Although the monsoon is identified by the reversal of wind pattern, its importance lies in the end product i.e., rain that occurs. Every monsoon is unique by itself as far as its onset, its further progression and spatial and temporal distribution of rainfall are concerned, even though the broad circulation features are similar in each case. This may be due to variation in the complex interactions taking place between different scales of motion, every year. Besides the intraseasonal variability, monsoon also exhibits interannual variability in its rainfall.

One of the interesting features associated with the South West monsoon activity it its intraseasonal oscillations which manifest themselves as meridionally propagating low frequency modes with a periodicity of 30 to 40 days, and which are related to the active & break conditions of the monsoon[3,4]. Krishnamurti and Subramanyan[5] presented a mapping of 30 to 50 day filtered wind fields at 850hpa over an extensive region during the period 2nd May to 28th July 1979, and found a steady northward propagation of anomalies of troughs and ridges from the Equator to about 30 N with a meridional phase speed of 0.75 deg Lat/day. Similar results were obtained by Murakami[6], Yasunari[7], Ahlquist[8]. Another important feature which enhances the activity of these low frequency modes, near 15N, is the passage of westward moving, synoptic scale, transient systems resulting in the increased convective activity[9]. Since the monsoon is a phenomenon of the Tropics, where the convective precipitation dominates, the importance of convection during this season need not be emphasized. The important role played by the latent heat release during convective precipitation in acting as a forcing mechanism of large scale tropical and extra tropical circulations is recognised by the numerical modellers long ago[10,11]. Apart from convection as such its diurnal varition which may be related to the transient part of the thermal forcing mentioned above, also needs to be taken into account for improving day-to-day forecasts. This information over oceanic areas, where conventional observations are scanty, is particulary useful. Fishing and oil Exploration operations can advantageuosly be planned if the diurnal variation of the hazardous weather associated with the convective clouds is known. In this context, the above information over oceans and coastal areas is of practical utility.

Most of the research work in the earlier years has been conducted using the observations from the conventional instruments. Since the deployment of these instruments over ocean areas is impractical, these regions remained as vast data gap regions. It is in this context, that the observations from space borne vehicles, such as satellites, assume importance in undertaking studies of the monsoon. The advantage with the satellite observations, specifically those from the Geostationary satellites, arise from the fact that the observations can be made over a larger area, at any time, at frequent intervals. Since clouds are visible manifestation of atmospheric circulation, cloud observation over a larger area through satellites form an excellent data set. By monitoring the changes in the cloud pattern with time the changes in the circulation pattern that occur as a result of the interaction between different scales of motion can be inferred. The quantitative information such as cloud motion vector (CMV), Quantitative precipitation estimates (QPE), Outgoing longwave radiation (OLR), and Sea surface Temperature (SST), derived from the satellite data can advantageously be used as parameters to study various facets of monsoon.

In this paper, it is proposed to present some aspects of the monsoon using cloud top temperature, as derived from IR channel data, of INSAT & GMS satellites.

Data

The Indian National Satellite (INSAT) is a multipurpose geostationary satellite which has been designed to cater to three different fields i.e., Television, Communication and Meteorology. The meteorological payload is a Very High Resolution Radiometer (VHRR) operating in two channels (*a*) Visible (0.55μm to 0.7 μm) and (*b*) Infrared 10.5μm to 12.5μm) with resolutions of 2.75km and 11km respectively at sub-satellite point. The IR sensor can discriminate temperature differences of 1°K or less. During normal operations ingests are processed every three hours. In the INSAT-1 series 4 satellites were launched, the last being INSAT-1D positioned at 83 deg E over the Equator. The first satellite in INSAT-2 series was launched in July 1992 and is positioned at 74E. The INSAT-2 series are similar to those of INSAT-1 series, with improved resolutions of 2km and 8km in the Visible and Infrared channels respectively.

The GMS satellite is a spinning satellite, unlike Insat which is three three axis stabilised, meant only for meteorological purposes. The Visible and Infrared Spin Scan Radiometer (VISSR) is a Visible and IR radiometer that provides imagery data. The VISSR is used to obtain Visible and IR images of the Earth and its cloud cover using an optical telescope and detector system. A scan of the Full Earth disk can be accomplished every 30 minutes using the 100 rpm spinning motion of the satellite for the west to east scan and the north to south scanning is done by stepping the scan mirror. The band width in the visible is 0.5 to 0.75 μm while it is 10.5 to 12.5μm in the IR with resolutions of 1.25km and 5km respectively. The Satellite is positioned at 140E over the Equator.

The author had the opportunity to work in a collaborative project in the Meteorological Research Institute, Tsukuba (Japan), which gave him the access to the data of GMS satellite, whereby the merging of the data of INSAT & GMS satellites was possible.

Three hourly average cloud top temperatures (CTT) over every 1 × 1 deg lat/long grids for the Monsoon season of 1991 over the area extending from 40E to 120E and 60N to 60S were prepared using the INSAT-1D radiance data. Similar data set for the GMS satellite were prepared for the area 80E to 160W and 60N to 60S, for the same period. These two data sets were merged after calibrating one with respect to the other over the common area equidistant from the two satellite positions (110E-115E). INSAT data were corrected according to the regression relationship derived, so that the data sets were compatible with each other. INSAT data were used west of 110E and east of 115E GMS data were used. Between 110E and 115E, the two data sets were gradually merged to avoid discontinuities. This composite data set has been used in the current study. Mean temperature of each 1 deg square mesh represents the mean height of the clouds in that area, at that hour except in cloudless areas where it indicates the surface radiative temperature. Since the information about the cloud height is derived from the temperature data, mean cloud height means the average of the tops of the clouds prevailing in the given area and does not refer to the average of the base and top of the clouds. Though this gives an average height of the could tops, in a given area, individual clouds may extend to still higher levels. As the mean temperature gives a good index of the convective activity in any square mesh, this parameter is chosen to discuss the cloud patterns in the following sections.

Earlier Studies on Monsoon using INSAT Data

Different aspects of monsoon were studied using satellite imageries, as such, and the derived data from the satellite radiance values. In a study of cloud patterns associated with monsoon depressions, Prasad *et al.*[12] found that a long zonal cloud band comprising deep layer cloud clusters is associated with weak

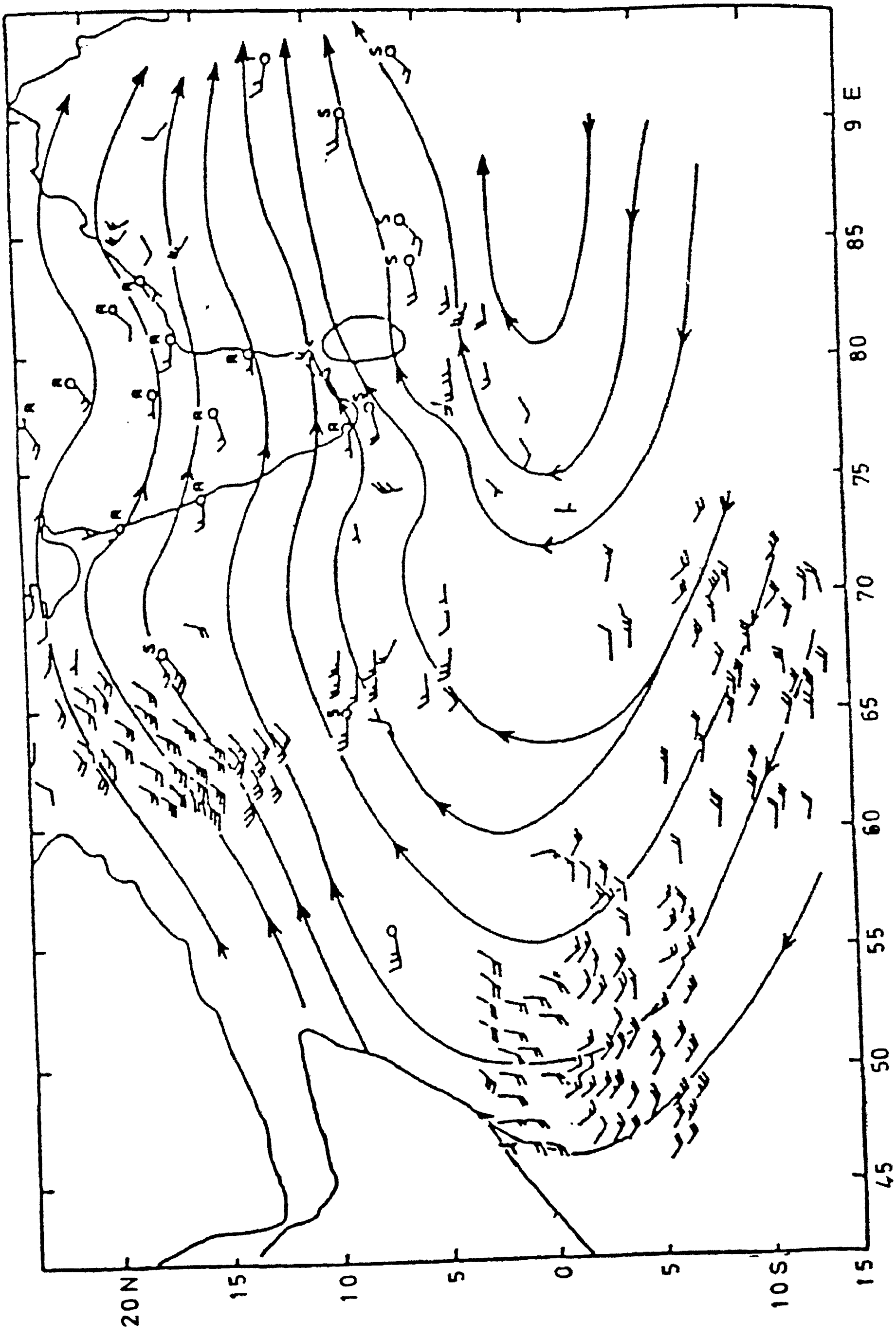


Fig. 1 06 UTC Low Level Cloud Motion Vectors for 30-May-1990 depressions or early stages of its evolution while in the case of an intense depression the clouds spiral towards the centre close to the edge of the dense overcast which has sharp edges and embedded overshooting tops. Kalsi and Halder[13], in a case study, studied the influence of westerly troughs on monsoon depressions, as revealed by the changes in the cloud pattern. Using enhanced IR imageries Mishra *et al.*[14], estimated rainfall in two heavy rainfall events of Aug '86 & '87, by adopting an approach similar to the life history technique of Scofield & Oliver[15], and concluded the results as encouraging. In a study of the movement of Inter Tropical Convergence Zone (ITCZ) during the onset phase of the monsoon with the help of satellite pictures Mandal and Bahuguna[16] found that the day-to-day fluctuations of ITCZ are not uniform.

Yadav and Kelkar[17] studied the onset of monsoon of 1988 using the Cloud Motion Vectors (CMV) indicating the importance of such winds in monitoring the onset of monsson. Typical flow pattern at the time of onset was given in Fig. 1. Sant Prasad *et al.*[18] presented a mean flow pattern and vergence fields derived, using CMVs over oceans during the monsoon season. Rao *et al.*[18] estimated the large scale precipitation and Outgoing Long Wave Radiation (OLR), using INSAT radiance data. Kelkar & Rao[20] studied the interannual variability of the monsoon rainfall for the years 1986, 87 & 88, using the rainfall estimates from satellite data. They presented the mean rainfall maps for the monsoon months and the anamolies from the mean for the individual months in each year. The anamoly maps clearly brought out the drought of 1987, with large negative anamolies. Making use of fractional cloud data (defined as the ratio of pixels colder than a given threshold temperature to the total pixels in a grid box of 2.5 Lat/Long) Rao *et al.*[21] studied the low frequency oscillations during the monsoon period of 1986 & 1987. Their time-latitudinal cross section indicated a clear northward progression of cloudiness from equator to 30N with a phase speed of 0.9° lat/day and with a periodicity of 30 days. They also subjected the data for spectral analysis for each block of 5° lat in width and 10° long in length and found two prominent modes one near about 30 days and the other about 10-15 days, the latter perhaps be pertaining to the synoptic scale transient systems indicated by Murakami *et al.*[9]. Rao and Rao[22], using the 3-hourly fractional cloudiness data for the years 1987-89, studied the diurnal variation of convective clouds over the area extending from 35N to 25S and 40E to 100E. The study indicated that fractional cloudiness is maximum over the north Bay of Bengal at 06UTC and it migrates westwards towards land upto 18 UTC and swings back to its original position by morning again. This East-West movement is not seen in convective cloudiness associated with the Southern Hemisphere Equatorial Trough (SHET). The westeard migration is more in August than in any other month. The OLR which can be used as a proxy measurement for convective activity is another important parameter. Kelkar, Rao & Sant Prasad[22] using 3-hourly OLR data, derived from INSAT-1B IR data for the years 1987 to 1989, studied the diurnal variation of OLR in all the four seasons of the year. They concluded that the OLR is not only low during the monsoon season but the hour of occurrence of minimum OLR over Indian subcontinent is also shifted by 6hrs from that in other seasons.

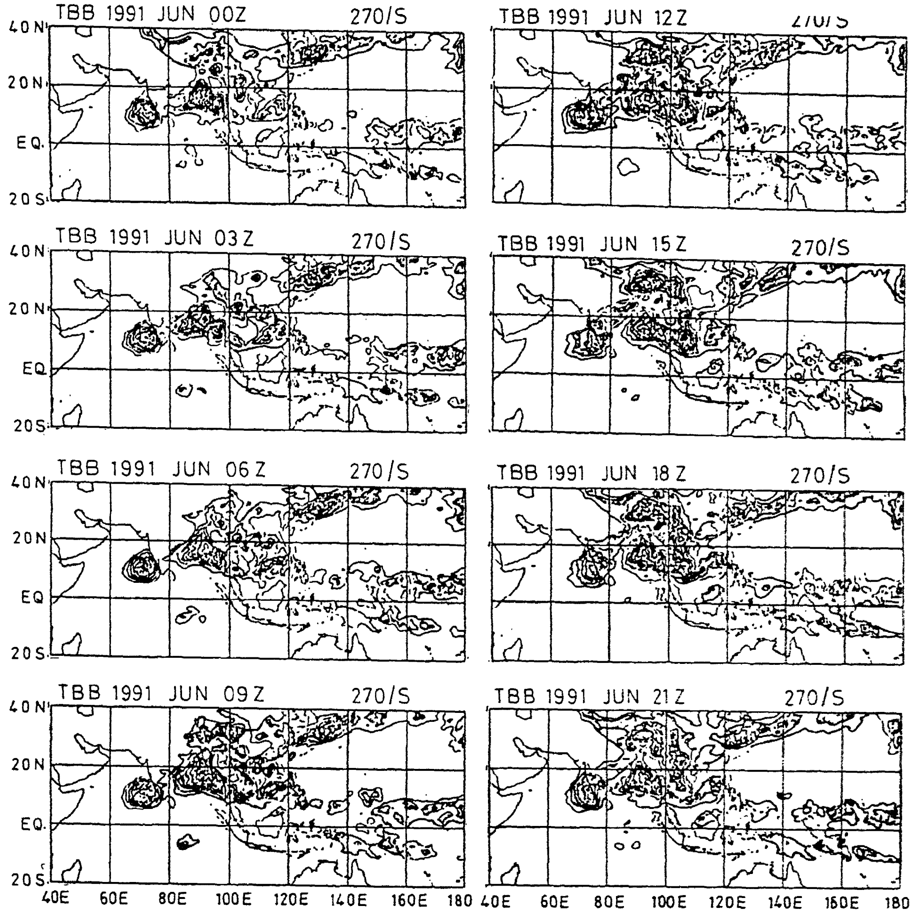


Fig 2(*a*) Horizontal Distribution of CTT for June, 1991

Results and Discussion

Mean CTT Patterns

The three hourly CTT patterns for the months of June to September are presented in Figs. 2(*a* to *d*) for the area 40E to 180E and 40N to 40S and the isopleths are drawn at 5K interval starting from 270K in the decreasing order.

June: In the month June, when the onset of monsoon has started, the axis of low CTT runs along 12N from 70E to 180E with a break from 120E to 150E, roughly indicating the position of ITCZ. This indicates that the convection is not uniform throught ITCZ, at least in the beginning stages of the monsoon. The SHET is not well defined, west of 140E during this month. The lowest CTTs are found only over SE Arabian Sea and over central and north Bay of Bengal indicating strong convection over these areas. The eastern protion of the ITCZ over East pacific Ocean is at a lower latitude than its western end.

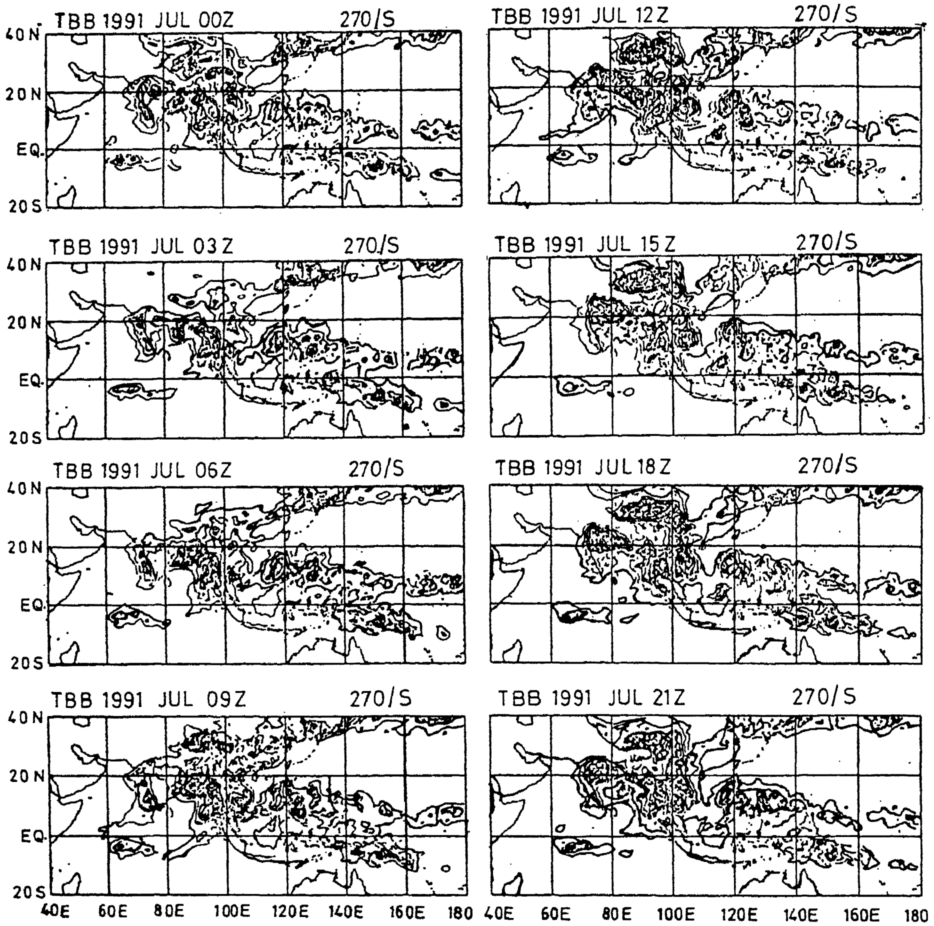


Fig 2 (*b*) Same as Fig 2(*a*) for July, 1991

July: In this month, when the monsoon advances over most parts of he countary, the convective activity in the ITCZ is more or less continuous. It is interesting to note that the eastern protion of the trough remains in the same position as in the month of June, over East Pacific while the western portion over India and neighbourhood shows northward progression. The SHET also is better defined over Indian longitudes than in the previous month. Another interesting feature is the east-west migration of convection with time over Indian region near head Bay which confirm the findings or Rao and Rao[22]. This indicates that the night time precipitation over central India more than that during day time. The rainfall in this region is produced mainly in squall lines or other long lived disturbances. The meso scale convective systems (which are common in the monsoon season) usually originate in the afternoon at about the time of maximum heating. Because of their relatively long life time their growth continues through the evening and their peak development and their maximum rain production is not reached until somewhat before midnight. The continued

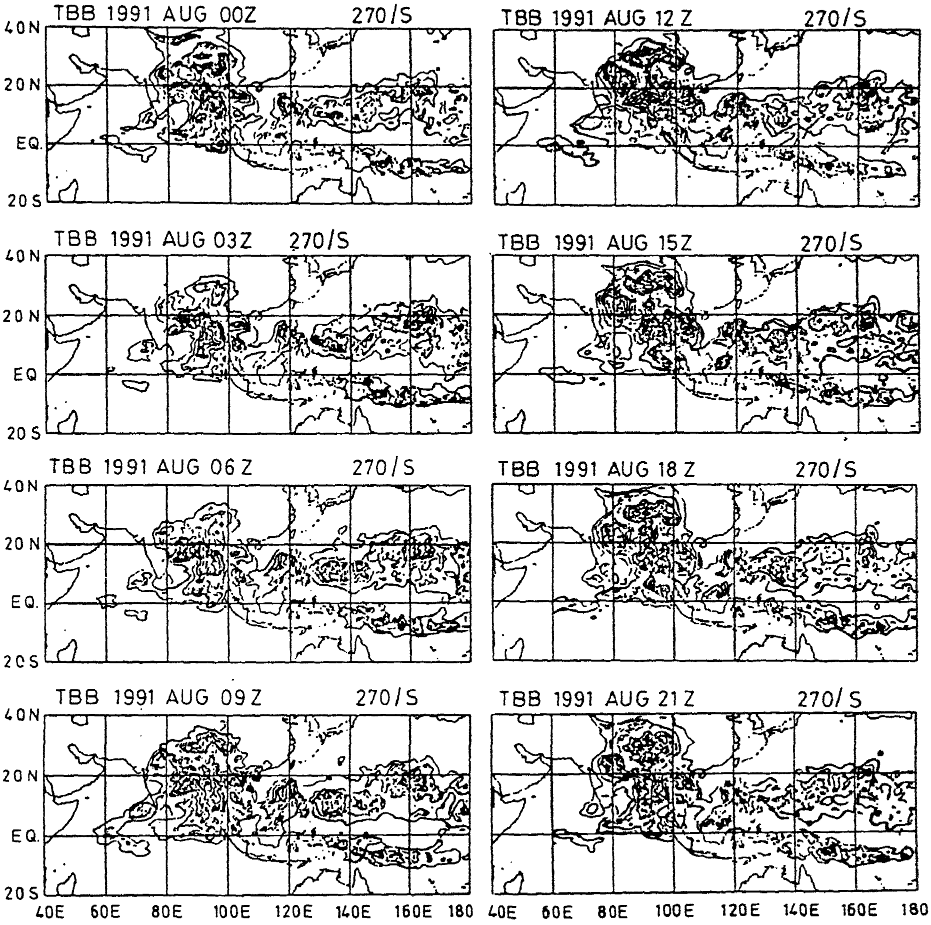


Fig 2 (*c*) Same as Fig 2(*a*) for August, 1991

spread of the cirrus shield delays the hour of maximum extent of bright or convective cloud until after midnight. This east-west oscillation is absent over the open ocean areas.

August: In August (Fig. 2*c*) the axis of lowest CTTs occupies a more northern position over all the longitudes including East Pacific. The CTTs are the lowest in this month indicating increased convective activity, compared to the previous two months. This is the month when the monsoon activity is at its peak. Another point of interest is the decrease of convection over SE Arabian Sea and the position of 270K isopleth over India which excludes the major portion of the west coast completely, where heavy rain occurs during this month. This clearly brings out the importance of orographic lifting in the occurrence of heavy rain in this region, from a comparatively warm top clouds. The east-west oriented low CTTs over east and central Pacific are due to the westward moving Typhoons during this year.

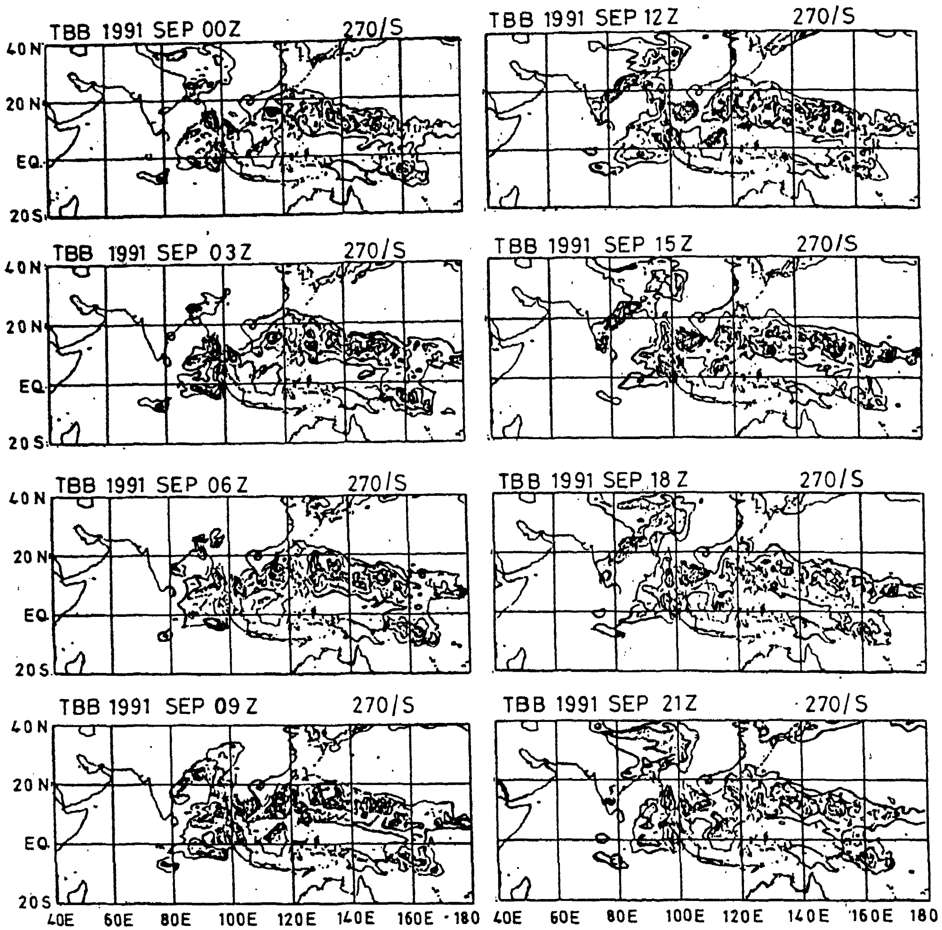


Fig 2 (*d*) Same as Fig 2(*a*) for September, 1991

September: This is the month in which monsoon withdrawal from the Indian subcontient starts. The CTT pattern (Fig. 2*d*) reflects this aspect very well by the absence of low CTTs over the country. Low CTTs are observed mostly east of 100E only. The activity over central and east Pacific continues. SHET moves south ward. The whole pattern shifts southeastward from that of August.

Two features which are commonly observed in all the four months are: (*a*) absence of convection over Sri Lanka and adjoining southeast peninsular India as indicated by the warm CTTs in all the months. (This region experiences minimum rainfall during this season); (*b*) Cloud top temperatures more than 270K in all the four months over Western Arabian Sea west of 60E.

Diurnal Variation of CTT

The importance of the diurnal variation of the convection is already mentioned in the earlier pages. Since the mean cloud top temperatures can be related to the mean height of the clouds, through the temperature-height relationship, the variation of CTT would give us a fair idea of the variation of the cloud height also. The lower the temperature the higher is the convective cloud top.

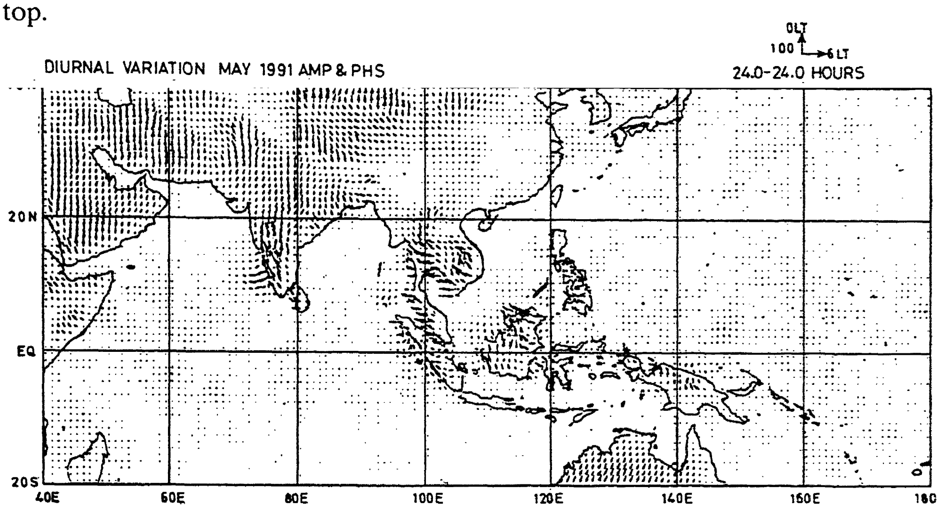


Fig 3 (*a*) Diurnal Variations of CTT for May, 1991

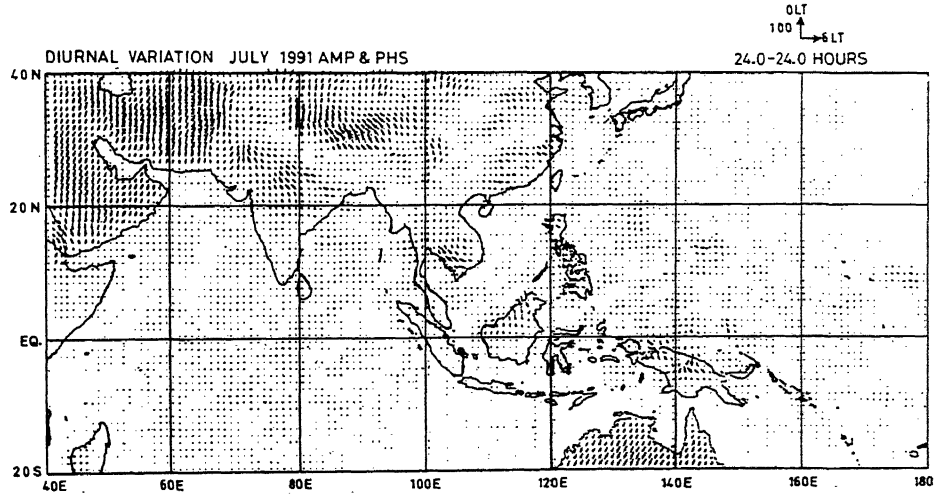


Fig 3 (*b*) Same as Fig 3(*a*) for July, 1991

The daily values of 3 hourly mean temperatures over 1 deg square mesh are averaged for each month. Using these values a fourier analysis was performed and the diagrams indicating the amplitude and phase are prepared. In these diagrams the length of the arrow is proportional to the amplitude. The northward pointing arrow pertains to 00hrs LT, the eastward arrow pertains to 06hrs LT, the southward arrow indicates 12hrs LT and the westward arrow indicates 18hrs LT. The direction of the arrow indicates the LT when CTT value at ath point becomes the lowest (i.e., heighest tops). Such diagrams are prepared for the months May to September 1991. The amplitude-phase diagrams for the months of May & July only are presented as representative months for premonsoon and monsoon seasons. Diurnal variations over different regions during all the monsoon months are presented below (Figs 3*a* &*b*).

South Arabian Sea: The amplitude is very small in general over this region becoming the smallest in the month of September. The lowest values occur around 19 hrs LT in the month of May while they occur around 16 hrs LT during the monsoon months. Variations are similar over equatorial Indian Ocean.

Central Bay of Bengal: (10N-15N, 80E-90E). Very little variation of CTT is noticed during May. Slight increase in the amplitude occurs in the monsoon season with the lowest temperatures tending to occur near 13hrs LT during the month August.

Monsoon Region: (22N-27N, 80E-90E). Over this region the preferred time of lowest CTT is 23hrs LT during May while it shifts 18hrs LT as the season advances with a progressive decrease in the amplitude becoming the lowest in the month of August.

Pacific Ocean: The diurnal variation of CTT is vary small over the Pacific Ocean and the minimum CTT tending to occur near 17hrs LT.

In general, the amplitude is large over the land areas particularly over pensular India, Thailand and adjoining Burma and with a sharp decrease during the monsoon season. The time of occurrence of lowest CTT is around 21hrs LT over all these regions in all the months.

Very little diurnal variation of CTTs occur over South China Sea and Japanese Islands.

Low Frequency Oscillations

Ever since the findings of Murakami[24], about the northward moving low frequency oscillations during monsoon season, these intraseasonal oscillations with a periodicity of 30-40 days had become a subject of intense research during the last decade. This 30-40 day oscillations appear to be present in many geophysical observations[5,25-29].

Since the fluctuations in the monsoon rainfall are directly related to those in the cloudiness, the data from the satellite can provide immediate and invaluable information on the large scale monsoon circulation system. The study of such oscillations using CTTs would reflect the fluctuations in the cloudiness. Earlier studies of these low frequency oscillations over the Indian monsoon region were carried out with twice daily values of the polar orbiting satellites. Owing to their orbital specification, analyses using these data are limited to some extent in time and space resolutions. In this respect, the data from the geostationary satellite which are more frequent and yield a better sampling for daily values would be better suited for such studies.

Using 3 hourly mean CTT values over each 1 deg square mesh, a parabolic curved line was fitted to seasonal variation. The objective is to remove the seasonal trend from the data. Anomalies were calculated as deviations from this curve. The data were then subjected to a band pass filter of 30-50 days. Time latitude cross sections and time longitude cross sections were prepared for these filtered anomalies. Time-latitude cross sections were prepared between

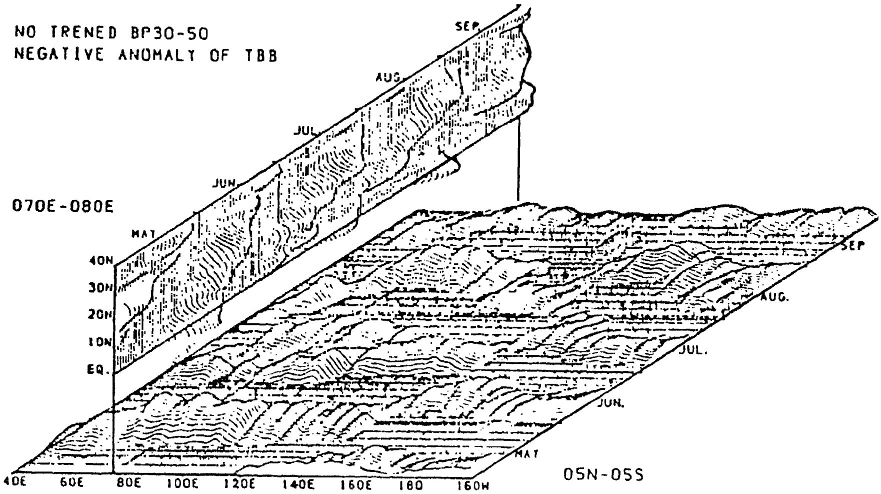


Fig 4 Time-longitude section for 5N-5S and time-latitude section for 70E-80E on the basis of 30-50 day band pass filtered anomaly. Only negative anomalics are depicted.

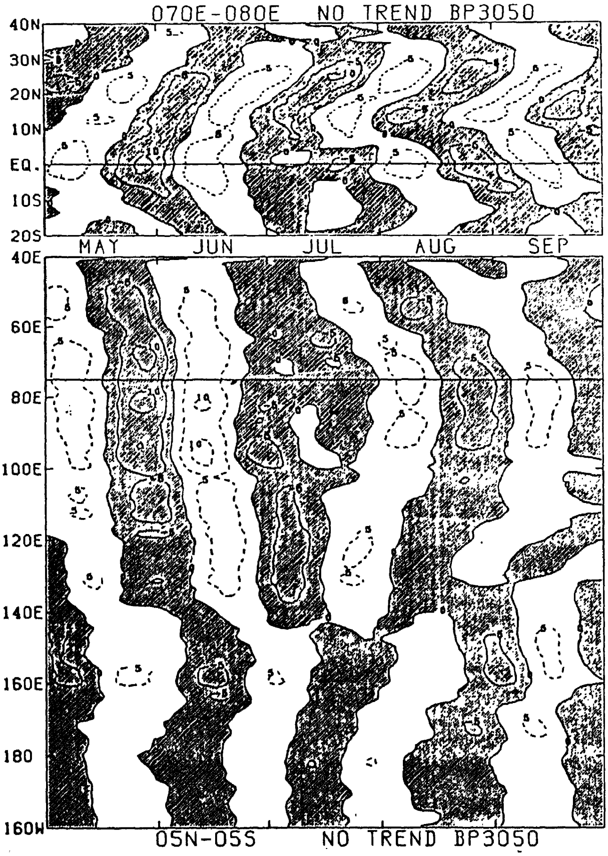


Fig 5 (*a*) Time-Latitude (70E-80E) and Time-Longitude (5N-5S) cross sections of temperature anamolies. Negative values are shaded. Isopleths at 5K interval.

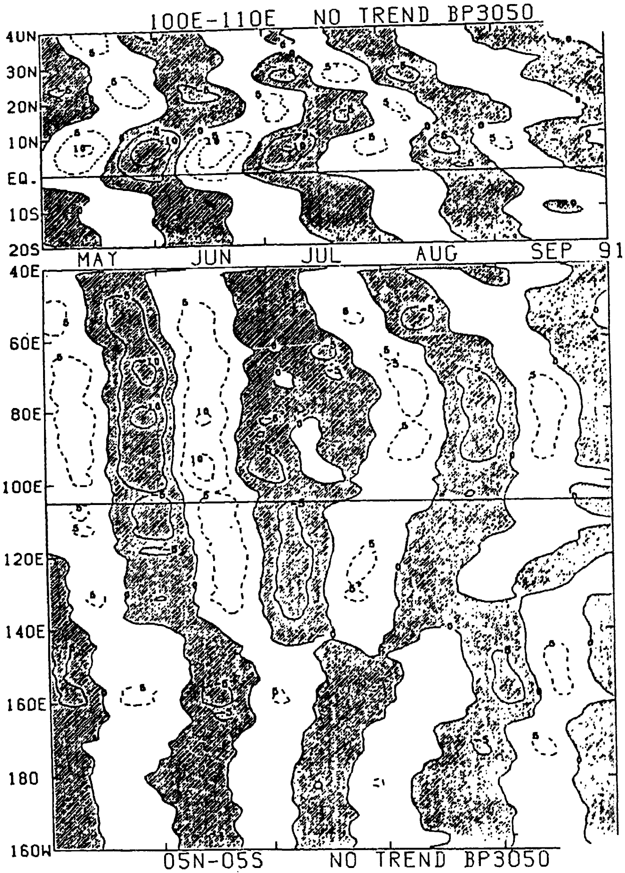


Fig 5 (*b*) Time-Latitude (100E-110E) and Time-Longitude (5N-5S) cross sections of temperature anamolies. Negative values are shaded. Isopleths at 5K interval.

50E-140E at each 10 deg longitude width (i.e., 50E-60E, 60E-70E etc.). Time-Longitude cross section between 5N-5S was also prepared. In Fig. 4, a composited view of time-latitude section between 70E-80E and a time-longitude section between 5N-5S is shown. Only negative anomalies are depicted. It is seen from this diagram that the maximum cloud zone (MCZ as indicated by the lowest anomalies), start from the equator and move northward upto 30N and disappear later. This sequence is repeated at an interval of about 40 days during the period May to September. It is also noticed that with each northward movement of cloud zone there is an associated eastward migration of equatorial cloud system as depicted in the time-longitude diagram. It is also observed that the northward migration is not only more prominent over Indian longitudes but also migrates to northernmost latitudes i.e., 30N while further eastward (east of 100E), the northward migration is restricted upto 20N only. Fig 5(*a&b*) give the time-latitude cross section between 70E-80E and 100E-110E along with time-longitude cross section of 5N-5S which clearly brings out the above observation. In this diagram, the negative anomalies are shaded and the positive anomalies > 5 K are indicated by the dotted curve.

It is interesting to note that the anomalies travel northward with an average phase speed of 1 deg Lat/day, a value comparable with that of Rao *et al.*[20]. Another cloud band (CTT minimum) starts at the equator as soon as the earlier one reaches the northernmost latitude. These events repeat with an average periodicity of 40 days.

Transient Systems

Another important feature which adds to the complexities of interaction between different scales of motion during the monsoon season, is the passage of westward moving synoptic scale transient systems over the Indian longitudes. Whenever a phase lock occurs between these systems and the northeard moving low frequency osillations, the convective activity in the monsoon trough is enhanced and the 30-40 day mode becomes prominent[9]. Rao *et al.*[20] investigated this aspect using fractional cloudiness of the two monsoons of 1986 and 1987. They prepared time-longitude cross sections for each 5 deg latitude zone between 5S-30N. They found that the movement of cloud maxima are towards east near the equator while the westward moving cloud maxima were observed

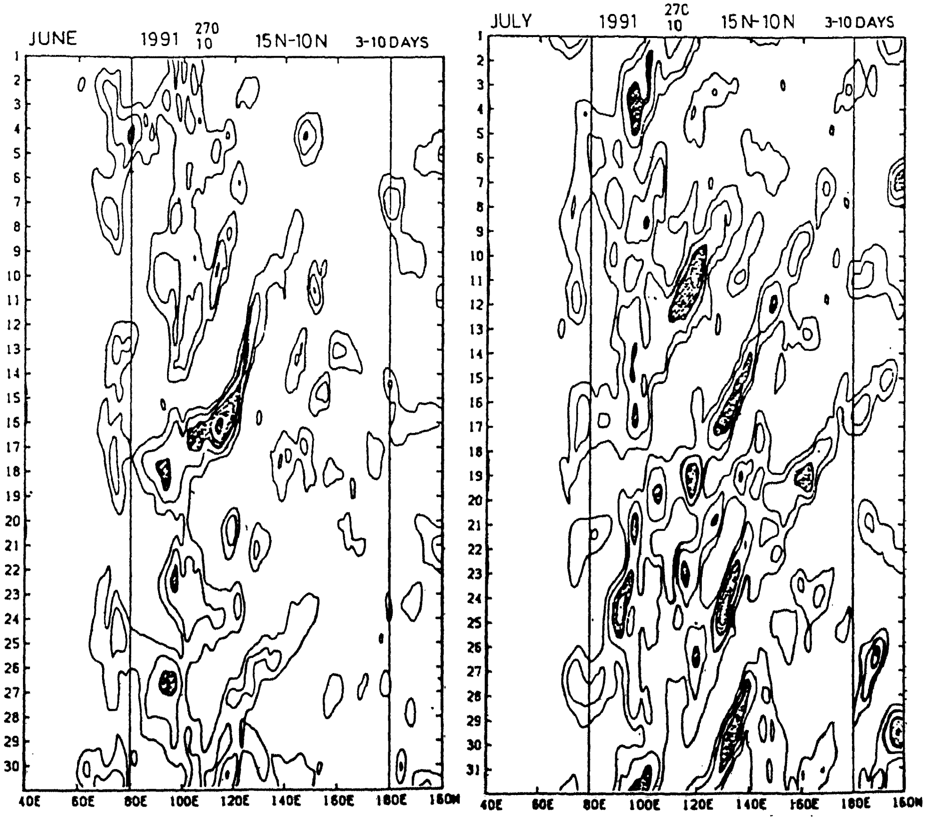


Fig 6 (*a*) Time-Longitude cross section between 10N-15N for June & July for filtered CTT data. Isopleths at 10K interval north of 5N with maximum activity between 10N-15N. They also observed the transient system activity is much supressed during 1987 compared to that of 1986 and concluded that this may be one of the contributory factors for below normal monsoon activity in 1987. These observations indicate that the westward moving transient systems play no unimportant role in keeping up the monsoon to be active.

In the cross sections presented by Rao *et al.*[21] the eastern boundary is 100E and most of the transient systems could be seen entering the area and hence their origin could not be traced, which may be far eastward. The compositive data set of INSAT & GMS satellites was very useful to examine this aspect.

Time-Longitude Cross Section: The 3 hourly mean CCT data over each 1 deg. square mesh over the area extending from 40E-160W were used in the preparation of these diagrams. Cross sections were prepared between 5S-30N for each 5 deg. latitude zone after applying 3-10 day band pass filter to this data. Time is depicted on the ordinate and longitudes on the abscissa. Isopleths are drawn at an interval of 10 deg. starting from 270 K and lower. Such cross sections are prepared for each month May to September. In Fig. 6(*a*)&(*b*) the cross sections for June, July, August & September were depicted for the latit-

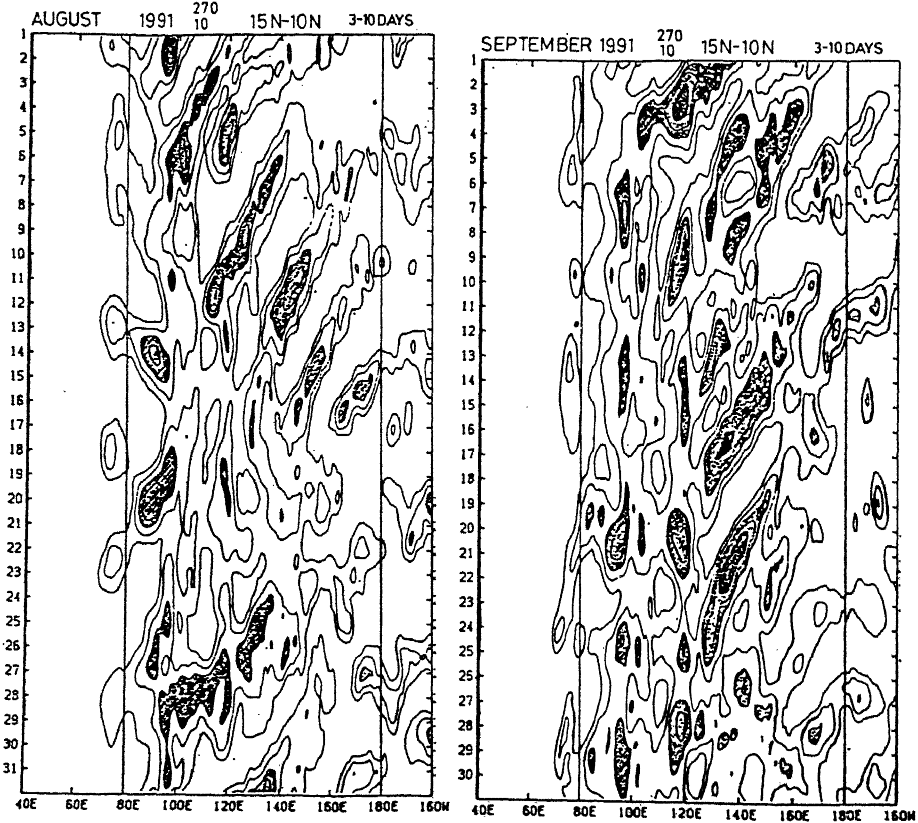


Fig 6 (*b* Time-Longitude cross section between 10N-15N for August 5 September. For filtered CTT data. Isopleths at 10K interval.

udinal belt 10N-15N only. The movement of the lowest CTT values indicate the movement of the convective activity associated with the transient systems. The westward movement is prominent.

From the examination of all the cross sections some general observations can be made:

(*a*) Near the equatorial region the cloud movements as indicated by those of CTT minima, are towards east while westward movement is noticed in the latitudes north of 5N.

(*b*) The number of such systems is maximum in the month of July

(*c*) The average life period of these systems is about 12 days while they last longer in July than in the other months.

(*d*) The preferred longitude of origin of these systems is between 150E-160E, though some of them (a few in number) could be seen to have originated at 160W. Most of these systems, having originated in the eastern pacific travel westwards and enter Bay of Bengal and lose their identity west of 80E.

(*e*) They travel with an average speed of 5 deg. longitude per day and these events repeat at an interval of 9 days.

(*f*) Such organised movement is absent in the month of May. This activity starts in the second week of June and continues during the whole monsoon season.

Summary

The cloud top temperature data (CTT) averaged over every 1 deg square mesh can be employed as one of the parameters to undertake studies of the monsoon. The composited CTT data, using INSAT & GMS satellites have been very useful in the study of some features of the monsoon. This study, undertaken as one of the projects under Indo-Japanese collabarative Research on Asian Monsoon proves the utility of satellite data in the study of different facets of the Monsoon.

(*a*) The maps of the horizontal distribution of CTTs indicate that intense convection occurs in preferred pockets along the ITCZ and is not uniformly distributed all along its length. While the western portion of the ITCZ shows northward movement over India and adjoining regions during the first two months of the season, the eastern end over Pacific remains more or less in the same region and moves northward only during August & September.

(*b*) Even within ITCZ, over the Bay and adjoining land areas, the convective activity exhibits east-west migration during the day with maximum occurring over the Bay near noon and that over land near midnight. Such east-west migration is absent over the open ocean areas.

(*c*) The negative CTT anomalies indicate a progressive northward movement starting from the Equator when viewed on a meridional plane, over Indian longitudes with a periodicity of 40 days while the anomalies on a latitudinal plane near the Equator indicate an eastward movement. Each epoch of northward movement of cloudiness is observed to coincide with that of an eastward movement of cloudiness in the Equatorial zone. This mode reaches the northernmost latitude over Indian region only while it terminates near about 20N, over the region East of 100E.

(*d*) The time-longitude cross-section of mean CTTs indicate that the lowest values of CTT, being a measure of convective activity move westwards, in association with the transient systems. These systems are not only longer in duration but are maximum in number also, in the month of July in the latitudinal belt of 10-15N. This movement appears to be necessary for good monsoon activity.

Acknowledgements

The author wishes to express his grateful thanks to the Director General of Meteorology for having given the opportunity to work at MRI, Japan. He wishes to express his sincere thanks to the Govt. of Japan who have provided the travel support and per diem during his stay at Japan. Thanks ar also due, to Sh P N Khanna and Dr V R Rao who helped in the generation of data set from INSAT.

References

1 Y P Rao *Southwest Monsoon, Met. Monogr. Synopt. Met* No. 1 IMD publication (1976)
2 P K Das *Monsoons* WHO Publ (1986)
3 T Yasunari *J met Soc Japan* **58** (1980) 225-229
4 T Yasunari *J met Soc Japan* **59** (1981) 336-354
5 T N Krishnamurty and D Subramaniam *J atmos Sci* **39** (1982) 2088-2095
6 M Murakami *Proc Sem Progr trop Met GARP Sp Rep* No 44 (1985) II-6 to II-13
7 T Yasunari *Proc Sem Progr trop Met GARP Sp Rep* No. 44 (1985) II-3 to II-5
8 J Ahlquist *Proc Sem Progr trop Met Grap Sp Rep* No. 44 (1985) III-14 to III-20
9 T Murakami, T Nakazawa and J He *J met Soc Japan* **62** (1984) 440-468
10 A C Gill *Q J R met Soc* **106** (1980) 447-462
11 K M Weickmann, G R Lussky and J Kutzback *Mon Wea Rev* **113** (1985) 941-961
12 K Prasad, S R Kalsi and R K Datta *Mausam* **41** (1990) 365-370
13 S R Kalsi and S R Haldar *Mausam* **43** (1992) 59-64
14 D K Mishra, S R Kalsi and R K Jain *Mausam* **39** (1988) 383-392
15 R A Scofield and V J Oliver *Preprint Volume of Second Conference on Flash Floods* Boston, MA (1980) 115-122
16 G S Mandal and G K Bahuguna *Vayu Mandal* **17** (1987) Nos 1&2, 67-70
17 B R Yadav and R R Kelkar *Mausam* **40** (1989) 325-328
18 Sant Prasad, P N Khanna, A V R K Rao and R R Kelkar *Mausam* **41** (1990) 445-450
19 A V R K Rao, R R Kelkar and P A Arkin *Mausam* **40** (1989) 123-130
20 R R Kelkar and A V R K Rao *Mausam* **41**(1990) 183-188
21 A V R K Rao, A K Bohra and V Rajeswara Rao *Mausam* **41** (1990) 51-58
22 A V R K Rao and V R Rao *Mausam* **44** (1993) 127-134

23 R R Kelkar, A V R K Rao and Sant Pradad *Mausam* **44** (1993) 45-52
24 T Murakami *J met Soc Japan.* **54** (1976) 175-181
25 D R Sikka and S Gadgil *Mon Wea Rev* **108** (1980) 1840-1853
26 J R Anderson and R D Rosen *J atoms Sci* **40** (1983) 1584-1591
27 T Murakami, T Nakazawa and M Jinhai Univ Hawai Deptt Met UHMET 83-02 (1983) 133 pp
28 Narayan R Iyengar *Mausam* **38** (1987) 453-458
29 U S De and D V Vaidya *Mausam* **38** (1987) 395-400

STUDIES ON ATMOSPHERIC AEROSOLS AT A COASTAL STATION, TRIVANDRUM

B V KRISHNA MURTHY

Space Physics Laboratory, Vikram Sarabhai Space Centre,
Trivandrum-695 022, India

(*Received* 3 *June* 1993; *Accepted* 5 *July* 1993)

Studies on atmospheric aerosols are being carried out at the tropical coastal station, Trivandrum (8°33'N, 77°E). These are based on the observations by Laser radar and Multiwavelength Radiometer. In this paper, some of the major scientific results obtained from these studies are presented.

Key Words: Atmospheric Aerosols; Aerosol Extinction; Tropical Region

Introduction

Atmospheric aerosols are defined as particles in liquid or solid phase suspended in the atmosphere. These particles cover a wide range of sizes extending from about 10^{-9} m to 10^{-4} m. Aerosols play an important role in the radiation budget of the earth-atmosphere system, atmospheric chemistry and atmospheric electrical conductivity.

Aerosols scatter and absorb electromagnetic radiation from the sun and the earth depending on their size distribution, composition (determining the refractive index) and wavelength of the radiation. The relative amount of absorption and scatter is governed by the aerosol single scattering albedo (w_0), the scattering asymmetry factor (g), the optical depth (τ) and the albedo of the underlying surface. If the aerosols are weakly absorbing type ($w_0 \approx 1$), they predominantly scatter the radiation, increasing the amount of solar radiation backscattered to space. This reduces the energy input to the earth-atmosphere system. The magnitude of this increase in planetary albedo is a function of aerosol size distribution and total optical depth. The asymmetry factor 'g' at a given wavelength increases with increasing particle size for a fixed aerosol composition. This means, that smaller particles tend to increase the planetary albedo in comparison with the large particles for the same optical depth. Absorbing aerosols ($w_0 < 1$) tend to reduce the solar energy lost from the earth-atmosphere system by absorbing the solar radiation. Therefore, this tends to increase the energy input to the earth-atmosphere system.

At infrared wavelengths, since the aerosol particle sizes are, in general, smaller than the wavelength and also since most materials exhibit absorption at these wavelengths, the single scattering albedo tends to be small. Thus, aerosols can be treated mainly as absorbers at infrared wavelengths. It follows that the main effect of aerosols at infrared wavelengths is to increase the total emissivity of the atmospheric layer in which they reside.

The radiative effects of aerosols, especially with regard to aerosol inputs due to major volcanic eruptions, have been studied by a number of workers[1,2,3,4,5]. Following the eruption of El Chichon volcano in 1982, a warming of the stratosphere by about a few degrees K was predicted and observations supported this prediction[4]. A similar warming of stratosphere has also been observed following the eruption of Mt. Pinatubo Volcano[5] in 1991. The model prediction for surface temperatures (for El Chichon and Mt. Pinatubo eruptions) are in the range of -0.5 °K (cooling) and so far there appeares to be no unambiguous observational evidence linking surface temperature changes with these volcanic eruptions[6].

The presence of aerosols in the atmosphere can support heterogeneous chemistry depending upon the atmospheric species taking part. Heterogeneous chemistry involves chemical reactions occurring on particle surfaces and in aqueous solutions. In recent years, the role of heterogeneous chemistry has come to be recognized as very important and crucial in the Antarctic ozone hole phenomenon. Reactions involving polar stratospheric cloud ice particle surfaces in the chlorine chemistry are generally believed to offer explanation for the Antarctic ozone hole phenomenon.

The role of aerosols in electrical conductivity of the stratosphere has been well recognized. Ions can act as nucleation centres to form aerosols and aerosols, in turn, can affect the ionization through heterogeneous chemistry. The presence of aerosols leads to decrease in electrical mobility and hence, in corresponding decrease in conductivity.

In view of the importance of aerosols in atmospheric processes as briefly outlined above, an understanding of the physical characteristics of atmospheric aerosols and their spatial and temporal variations is essential. While there is considerable information on stratospheric aerosols, information on tropospheric aerosols is rather scanty, especially at tropical latitudes. A programme of research on Atmospheric aerosols was started in the Space Physics Laboratory, Trivandrum in the early eighties and it is a continuing programme of the Laboratory. Some of the main scientific results obtained from this programme are presented in the following.

Ruby Laser Radar

Using a high power pulsed Ruby laser radar (Ruby Lidar), studies on tropospheric and stratospheric aerosols have been carried out at Trivandrum. Lidars essentially provide altitude resolved information on aerosol extinction/backscatter coefficients. The ruby lidar at Trivandrum consists of a ruby laser transmitter at 694.3 nm wavelength with a pulse energy of ~ 1J in a pulse width of ~ 4μs. The laser beam is directed vertically up into the atmosphere. The backscattered laser signal by atmospheric molecules and aerosols is received using cassegrain type telescope consisting of a 12″ parabolic primary mirror and a 4″ hyperbolic secondary mirror. The received scattered signal is collimated using a suitable optical system and is passed through an interference filter whose centre wavelength of transmission is 694.3nm. The filtered signal is detected using a cooled photomultiplier and the detected signal is processed and recorded for analysis.

The backscatter signal power $p(h_1)$ from a height interval $\Delta h(=c\ \Delta t/2)$ centred at altitude h_1, is given by

$$p(h_1)=\frac{P_0 \Delta h}{2} A \frac{\beta(h_1)}{h^2} \exp\ -2 \int \alpha(h)\,\mathrm{d}h$$

where

P_0 = transmitted power,
c = speed of light,
Δt = pulse width,
h = altitude,
E = optical efficiency of the transmitter and receiver optics,
A = Effective area of the receiving mirror,
$\beta(h_1)$ = volume backscatter coefficient at h_1
and
α = extinction coefficient.

$\beta(h_1)$ is given by

$$\beta(h_1)=n_a(h_1)\sigma_a+n_g(h_1)\,\sigma_g, \qquad \dots (2)$$

where σ_a and σ_g are the aerosol and molecular backscattering cross sections respectively and n_a and n_g are the number densities of aerosols and air molecules respectively. Similarly α is given by

$$\alpha(h)=n_a(h)\rho_a+n_g(h)\rho_g \qquad \dots (3)$$

where ρ_a and ρ_g are the extinction cross sections of aerosols and air molecules respectively. The lidar equation (equation 1) can be written in differential form, as

$$\frac{\mathrm{d}s}{\mathrm{d}h}=\frac{1}{\beta}\frac{\mathrm{d}\beta}{\mathrm{d}h}-2\alpha, \qquad \dots (4)$$

where

$s=\ln[p(h)h^2]$.

Eq. (2) contains two unknowns, namely α and β. Assuming a functional form of relationship between α and β the above equation can be solved[8,9] for α (or β) Using Fernald's method[9] of solving equation (2) with appropriate functional relationship between α and β, lidar observations of backscatter signal at Trivandrum have been analysed to obtain altitude profiles of aerosol extinction[7] Parameswaran *et al.*[10] carried out a comparison of aerosol extinction profiles obtained at Trivandrum and those obtained from SAGE II (Stratospheric Aerosol and Gas Experiment) satellite observations. The comparison showed a fairly good agreement with regard to both stratospheric and tropospheric aerosol extinction profiles. The comparison also revealed an increase in stratospheric aerosol extinction with decrease in latitude even within the tropical region. Further, the tropospheric aerosol extinction profiles revealed significant day to day as well longitudinal variabilities.

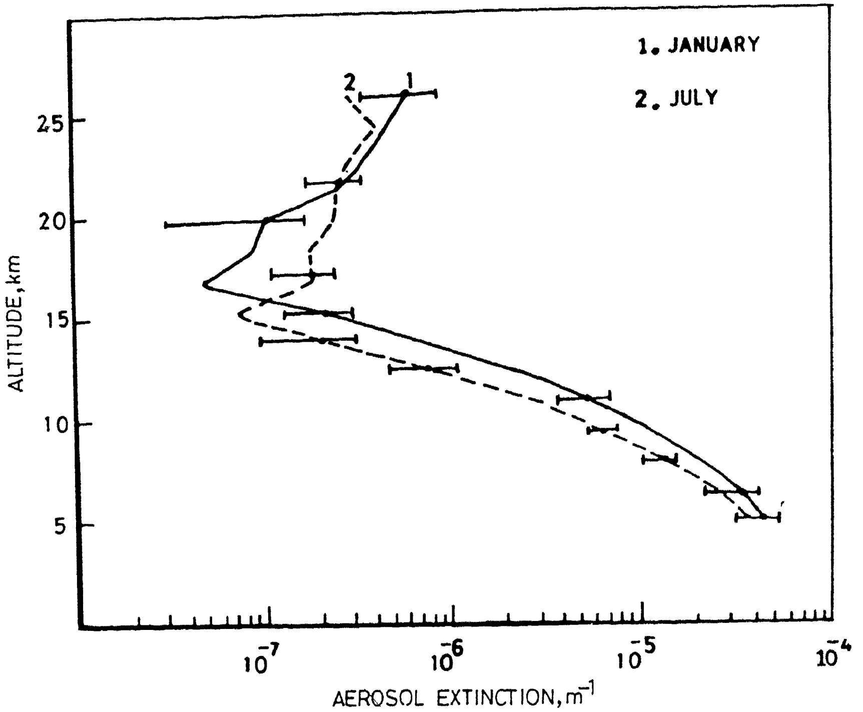


Fig 1 Monthly mean altitude profiles of aerosol extinction obtained from ruby lidar experiment at Trivandrum for (1) January 1987 and (2) July 1987

Tropospheric Aerosol Extinction

Parameswaran *et al.*[11] carried out a detailed study of the altitude profiles of tropospheric aerosol extinction and their temporal variations from lidar observations at Trivandrum over a period of about one and half years. They used the monthly mean profiles of aerosol extinction for this purpose. In Fig. 1 are shown the monthly mean profiles of aerosol extinction for January 1987 and July 1987 as typical examples. The lines parallel to the abscissa (in Fig. 1) show the standard deviations representing the day-to-day variations of the extinction. The altitude profiles show a smooth decrease with altitude up to about 17km (tropopause altitude) above which the extinction increases (in the stratosphere).

Making use of the monthly mean aerosol extinction profiles Parameswaran *et al.*[11]. estimated the (vertical) eddy diffusion coefficient at tropospheric altitudes. Considering steady conditions, the equilibrium aerosol number density in the middle and upper tropospheres can be represented by the continuity equation as

$$\frac{\partial n}{\partial t} = q - an - bn^2 - \frac{\partial}{\partial h}(wn) + \frac{\partial}{\partial h}\left(D\frac{\partial n}{\partial h}\right) = 0, \quad \dots (5)$$

where n is the aerosol number density,

q is the aerosol production rate,
D is the eddy diffusion coefficient,
b is the coagulation coefficient,
a is the rate of loss by washout process,
and
w_1 is the vertical velocity (gravitational sedimentation velocity).

In eq. (5), the aerosol production term, q is neglected as in the middle and upper tropospheres, the local production which is mainly by the gas to particle conversion process can be neglected in comparison with the other terms in the equation. In the aerosol size range of 0.02μm to 10μm which contributes to the backscatter/extinction at 694.3nm (ruby lidar operating wavelength) and at altitudes under consideration, the coagulation and gravitational sedimentation terms can be neglected. With these simplifications, the diffusion coefficient D can be obtained from eq. (5) as[11]

$$D \simeq \left(\frac{\partial n}{\partial h} \right. \qquad an\,\mathrm{d}h \tag{6}$$

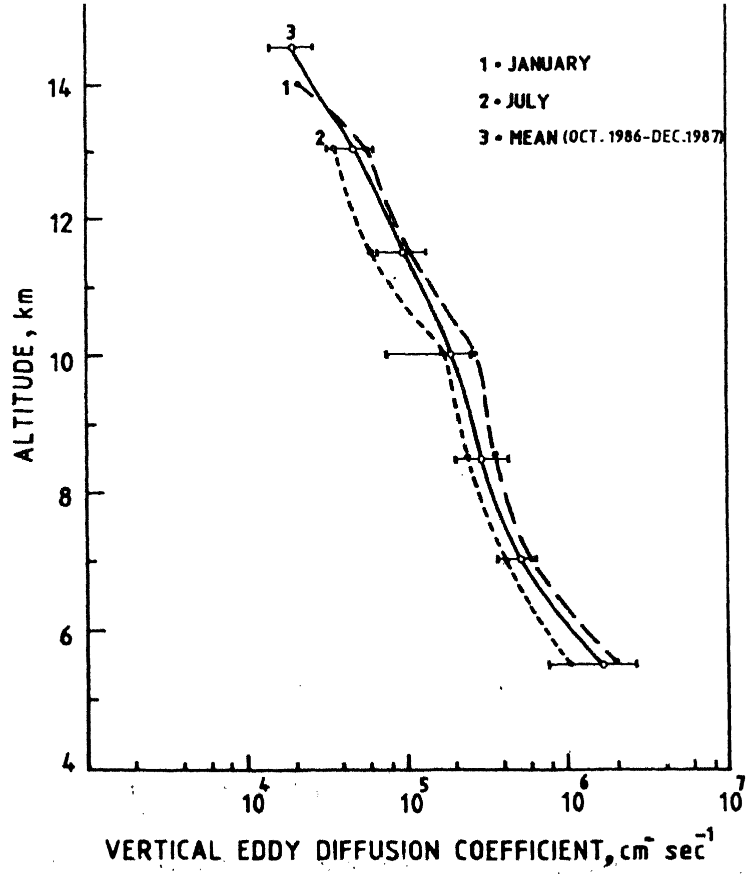


Fig 2 Mean altitude profile of vertical eddy diffusion coefficient for 1) January 1987, 2) July 1987 and 3) for the period October 1986 to December 1987

Adopting the model given by Turco *et al.*[12] for '*a*' Parameswaran *et al.*[11] estimated the vertical eddy diffusion coefficient '*D*' and their results are shown in Fig. 2 for the months January, 1987 and July 1987. The mean value of *D* (for the period of lidar observations from October 1986 to December 1987) is 4.8×10^5 cm^2 sec^{-1} at 5.5km altitude and decreases in nearly exponential manner with altitude. Parameswaran *et al.*[11] compared the diffusion coefficient profiles obtained by them with those obtained by earlier workers using data of Radon 222 gas and chemical releases and found good agreement. Thus, aerosol extinction or number density profiles can be used to estimate the vertical eddy diffusion coefficients in the middle and upper tropospheric regions.

Parameswaran *et al.*[11] studied the temporal variation of tropospheric aerosol extinction at different altitudes. Fig. 3 is a reproduction of their results. Shown in the same figure are also the integrated aerosol extinction from 4.8km to 16.8km and the monthly total rainfall at Trivandrum. The variations (with month) of aerosol extinction at altitudes $\leqslant$10.8km are quite similar whereas at higher altitudes, the variations are different from these especially in April-May,

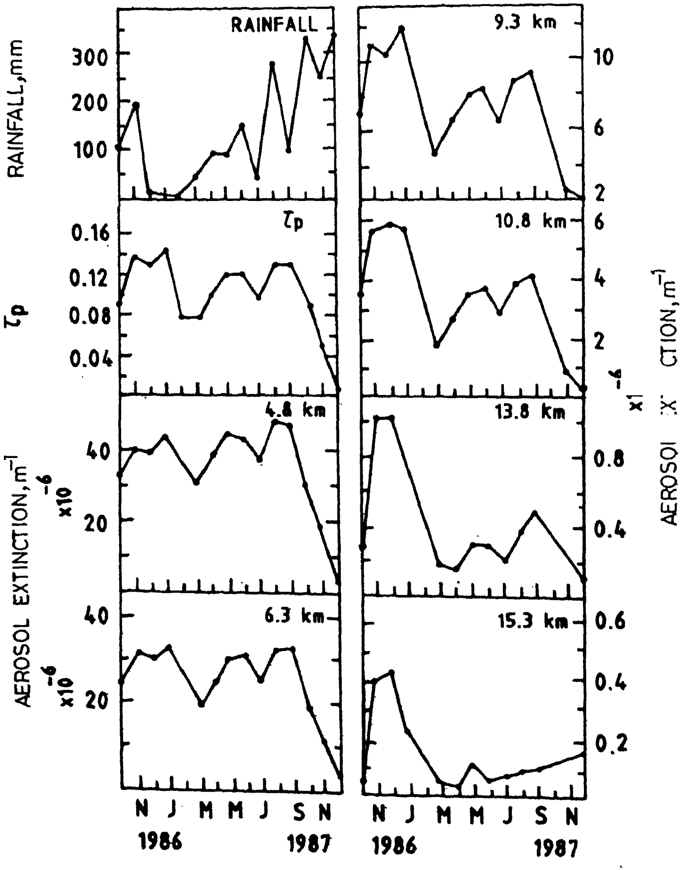


Fig 3 Temporal variation (*a*) of monthly mean aerosol extinction at different altitudes in the troposphere (*b*) τ_p, the total optical depth for the altitude range 3.8km to 16.8km and (*c*) monthly total rainfall

1987. There is a peak in the winter months of 1986-87 which is more prominent at higher altitudes. The sharp decrease in aerosol extinction at altitudes below 15.3km as well as in the integrated extinction during October-December 1987 can be attributed to strong wet removal of aerosols because of the heavy rainfall as indicated in the figure. This decrease (in extinction) is not seen at higher altitudes indicating that this loss process is not effective at these altitudes. The aerosol extinction peak during the summer/monsoon months can be attributed to *i*) an increase in production of aerosols from a possible increase of photochemical reactions due to volatile organic plant materials of plant origin, *ii*) an increase in tropospheric vertical mixing caused by high convective activity and *iii*) an increase in atmospheric water vapour content resulting in growth of aerosol particles. Further, at Trivandrum, strong surface winds during summer and monsoon months cause production of sea-spray aerosols due to sea-surface agitation[13]. These processes also could contribute to the peak in summer/monsoon months.

Stratospheric Aerosol Extinction

Stratospheric aerosols are mainly comprised of sulphuric acid-water solution droplets though ammonium sulphate is also present[14]. The sulphate particles are produced locally in the stratosphere mainly from precursor sulphur bearing vapours. Sulphur dioxide (SO_2) and carbonyl sulphide are the major precursor gases. SO_2 which is an important component of volcanic effluvia, is injected directly into the stratosphere during major volcanic eruptions. For instance, the eruption of Mt. Pinatubo volcano (Philippines) in June 1991 is estimated to have injected about 20 million tons of SO_2 into stratosphere which is about 3 times more than that due to the earlier volcanic eruption of El Chichon (Mexico) in March/April 1982. During volcanically quiescent periods, carbonyl sulphide (OCS) is the major precursor gas with carbon disulphide (CS_2) as a possible second source[16,17]. OCS originates mainly from biological and industrial sources and has very long tropospheric lifetime of several years[12]. In addition to OCS and CS_2, SO_2 from surface emission is transported to stratosphere during volcanically quiescent periods also. However, modelling studies[18,19] showed that OCS is the dominant sulphur source compared to CS_2 and SO_2 for stratospheric aerosols (during volcanically quiescent periods). The precursor gases are transported into stratosphere through eddy diffusion and direct injection which can occur during thunder cloud penetration, tropopause folding and vertical convection especially in the tropical zone. Horizontal advection subsequently distributes the injected material zonally and meridionally. OCS breaks up into sulphur and carbon monoxide by photodecomposition due to absorption in the extreme ultra violet (200-240nm) in the stratosphere. The sulphur atoms then react rapidly with molecular oxygen to form SO which on oxidation becomes SO_2. SO_2 is transformed into H_2SO_4 mainly through reactions involving OH radical. The H_2SO_4 thus produced and H_2O vapours nuclease by heterogeneous heteromolecular and homogeneous heteromolecular processes. In the former process, growth of pre-existing particles takes place due to condensation of H_2SO_4 and H_2O vapours. In the later process, new particles are formed from gas phase consisting of H_2SO_4 and H_2O vapours and the newly formed particles further grow through condensation of these vapours. It is generally believed that the heterogeneous heteromolecular process is more efficient than the homogeneous heteromolecular process. However, recent modelling calculations in this regard[20] showed that at temperatures around -75°C which are not uncommon in lower stratosphere, the homogeneous heteromolecular process can be a significant contributor to the overall nucleation process. Thus, stratospheric aerosols can be expected to show a significant dependence on temperature, Krishna Murthy *et al.*[21] investigated this aspect in detail using the stratospheric aerosol extinction profiles obtained from lidar observations at Trivandrum.

In Fig. 4 (from Krishna Murthy *et al.*[21]) is shown the variation of the monthly mean integrated stratospheric aerosol extinction and the monthly mean tropopause temperature. The variations of these two parameters are opposite to each other indicating a strong negative correlation between the two. The aerosol extinction at different stratospheric altitudes also showed similar monthly mean variations. Krishna Murthy *et al.* estimated the correlation coefficients between these parameters as shown in Table I. In this table, the correlation coefficients significant above $P = 0.05$ level[22] are indicated by asterisk.

The correlation (direct) of extinction at different altitudes with tropopause temperature is negative and significant whereas with ambient temperature, it is not significant though negative. A decrease in tropopause temperature is asso-

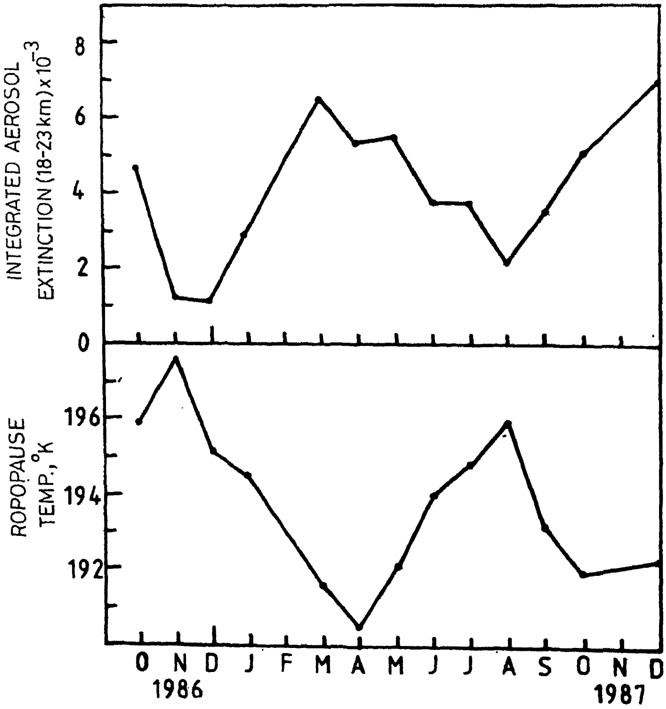


Fig 4 Temporal variations of monthly mean tropopause temperature and monthly mean integrated aerosol extinction (in Stratosphere)

Table I

Direct correlation coefficient (R) *between aerosol extinction* (α_a) *at different stratospheric altitudes, ambient temperature* (τ) *tropopause temperature* (τ_p), *and partial correlation coefficients* (r) *between* α_a *and* τ_p *eliminating effect of* τ *and between* α_a *and* τ *eliminating effect of* τ_p. *Correlation coefficients significant at* $p=0.05$ *level are indicated with superscript of asterick*

Altitude km	$R(\alpha_a, \tau)$	$R(\alpha_a, \tau_p$	$R(\tau, \tau_p$	$r(\alpha_a, \tau_p)$	$r(\alpha_a, \tau)$
18.3	−0.39	−0.65*	0.51	−0.57*	−0.09*
19.8	−0.44	−0.69*	0.40	−0.75*	−0.26*
21.3	−0.37	−0.72*	0.16	−0.72*	−0.36*
22.8	−0.17	−0.82*	−0.19	−0.88*	−0.57*

ciated with an increase in tropopause altitude which is indicative of stronger tropospheric convective activity[23]. This results in injection of significant tropospheric mass containing small size aerosol particles and aerosol precursor gases[24], OCS and SO_2. The injection of these constituents will lead to stratospheric aerosol nucleation and growth by condensation as discussed above and lower temperatures (ambient) are conducive for these processes. It is interesting to note that the partial correlation of aerosol extinction with ambient temperature eliminating the effect of tropopause temperature at 22.8km is significant whereas that at lower altitudes is not significant. These results clearly indicate that at lower stratospheric altitudes (<22km), aerosol extinction can mainly be attributed to direct particle injection across the tropopause whereas at higher altitudes, the stratospheric microphysical processes (heterogeneous heteromolecular and homogeneous heteromolecular processes) which are favoured by lower ambient temperatures become important. Further, Krishna Murthy *et al.*[21] showed that the observed increase in stratospheric extinction associated with lower temperature is consistent with the theoretical model calculations on stratospheric aerosols by Yue and Deepak[20]. These studies based on lidar observations at Trivandrum clearly brought out the effects of temperature on stratospheric aerosol extinction.

Studies on Total Aerosol Optical Depth

Total aerosol optical depth (τ_p) is very useful parameter in studies on temporal and spatial variations as well as radiative effects of aerosols. With a knowledge of τ_p and aerosol size distribution, the atmospheric path radiance and transmission coefficient can be obtained from approximate formulation of solution of radiative transfer equations assuming refractive indices[25]. As the chemical nature of different types of aerosols is known, refractive indices or at least the range of indices applicable can be inferred with a reasonable degree of confidence depending on the location. Study of aerosol total optical depth, size distribution, their temporal variations and long term trends would help in the understanding of the nature of aerosols and the relative contributions from natural and man-made sources. With this objective in view, study of total aerosol optical depths at different wavelengths in the range 400nm to 1025nm was initiated in November 1985 at Trivandrum. A multi-wavelength solar radiometer (MWR) has been set up in November 1985 at Trivandrum for this purpose[26] Similar radiometers have been set up later under the Indian Middle Atmosphere Programme, at Mysore, Visakhapatnam, Jodhpur and New Delhi. The operation of these radiometers is being continued as part of the Geosphere-Biosphere Programme.

The radiometer operates at nine wavelengths, namely 400nm, 450nm, 500nm, 590nm, 700nm, 750nm, 800nm, 935nm and 1025nm with a bandwidth of 5nm. An additional wavelength 500nm is incorporated in September 1987. The direct solar flux is measured as a function of solar zenith angle. Using the standard Langley method, the total atmospheric optical depth (τ) is obtained from these measurements. τ consists of three components namely, *i*) molecular (Rayleigh) scattering, *ii*) gaseous absorption and *iii*) aerosol scattering and absorption. The first component is estimated by using the relevant expression for Rayleigh scattering and air densities[26]. The second component which is mainly due to ozone at some of the wavelengths (590nm) is estimated using the relevant absorption cross sections and densities of the corresponding

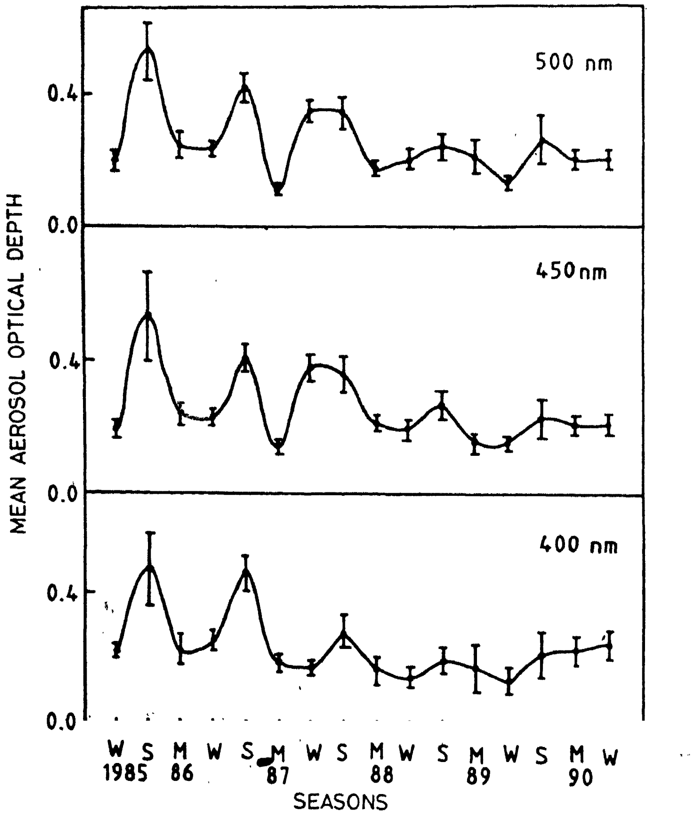


Fig 5 Temporal variations of seasonal mean aerosol optical depths at wavelengths 400nm, 450nm and 500nm. In the figure *W* indicates winter season comprising of December, January and February; *S* indicates summer season comprising of March, April and May months, *M* indicates monsoon season comprising of months from June to November

atmospheric constituents[26]. These two components thus estimated are then subtracted from τ to obtain total aerosol optical depth τ_p. The wavelength 935nm is used to estimate the total atmospheric water vapour content making using of the relatively strong water vapour absorption at this wavelength[13]. Altitude profiles of aerosol extinction obtained from lidar experiment[7] show that nearly 90% of the total aerosol extinction (optical depth) is contributed from the first 8 km and about 20% from the first 1km of the atmosphere from the ground. Thus, the total aerosol optical depths can mainly be attributed to aerosols in the altitude range 0-8km.

Figs 5, 6 and 7 show the variations of seasonal mean aerosol optical depths (at Trivandrum) at different wavelengths. It is readily seen from the figures that τ_p attains its maximum in summer and broad minimum in monsoon-winter seasons. However, the summer maximum becomes weaker and less pronounced in the later years (of the period of observations). Further, the seasonal variations appear to be wavelength dependent, with regard to their maximum to minimum ratio in an year. The standard errors are, in general, greater in summer than in the other two seasons, indicating greater day to day variability.

The seasonal variations of aerosol optical depths are to be understood in terms of the variations in different aerosol production, loss and transport processes. The main aerosol production/growth and loss processes are *i*) gas-to-particle conversion due to photochemical activity *ii*) particle growth by condensation of atmospheric water vapour *iii*) generation by wind and *iv*) rainout/washout (or wet removal processes). The first three processes of production/growth are likely to be prominent during local summer season. In addition to these, other aerosol sources and sinks of regional importance must also be considered. As Trivandrum is a coastal station, a significant aerosol input of marine origin can be expected. The strong westerly winds as well as the heavy rainfall associated with monsoon[27,28] are important in contributing to aerosol optical depth at this station. The south west monsoon setting-in during the first week of June at the coast of Trivandrum brings significant changes in tropospheric circulation earlier than its onset[29,30]. Strong surface westerly winds set-in from April onwards. These westerly winds gain strength with the advance of summer and results in considerable agitation of sea surface which manifest as sea surf and white caps. Considerable amount of marine aerosols is produced due to sea surf (and white caps) whose size spectrum and abundance depend upon the wind speed[31]. These marine aerosols are brought over to coastal land by the westerlies. As already pointed out, the wet removal process (rainout/washout) is a very important process of tropospheric aerosol removal and its strength is mainly determined by the rainfall. It should also be noted that surface wind (westerly monsoon winds) and rainfall which are effective in production and loss of aerosols respectively, themselves are associated with each other. In order to determine the effects of these two parameters on aerosol optical depths, correlation coefficients between these parameters are obtained. For this purpose the seasonal mean surface zonal wind speed (measured at a location near the MWR) and rainfall (of Trivandrum) are used. The correlation coefficients are shown in Fig. 8 for different wavelengths. Along with the direct correlation (r) between aerosol optical depth (r_{12}) with rainfall and wind (r_{13}),

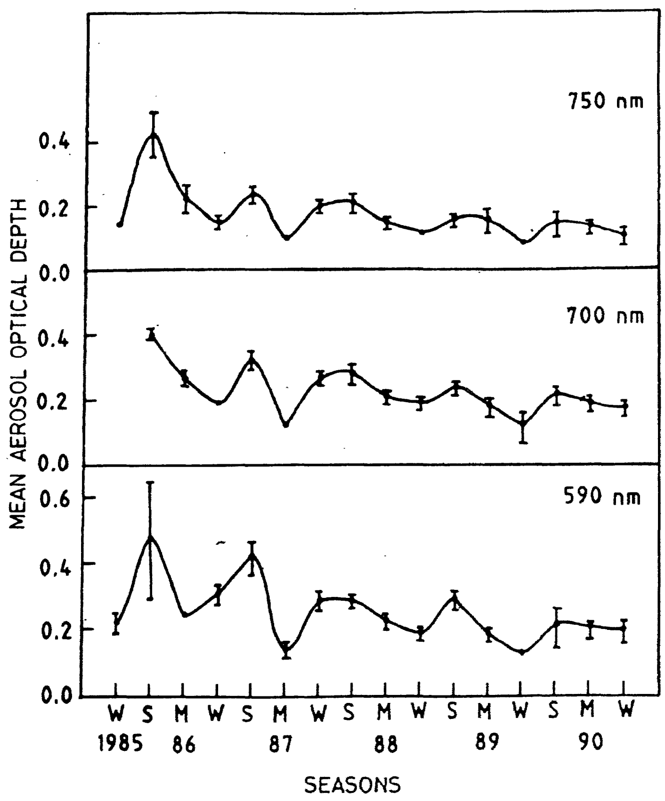


Fig 6 Same as Fig 5 except for 590nm, 700nm and 750nm

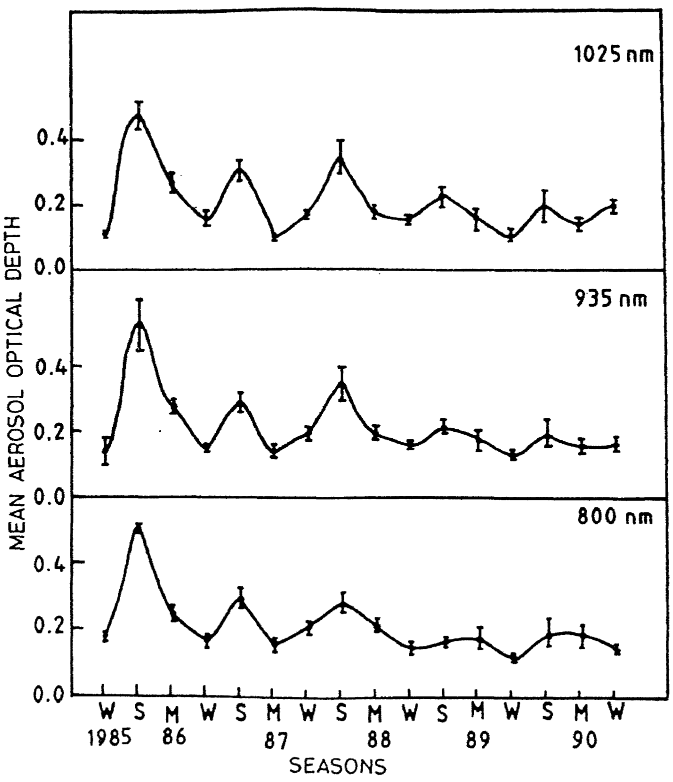


Fig 7 Same as Fig 5 except for 800nm, 935nm and 1025nm

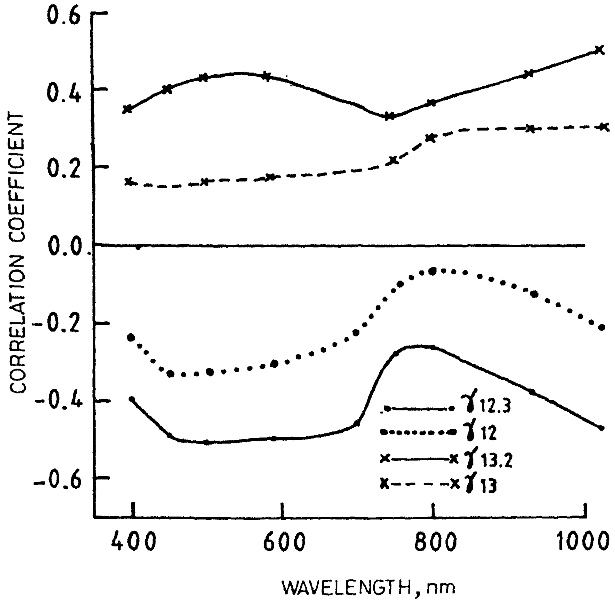


Fig. 8 Correlation coefficients (r) as function of wavelength. In the figure suffixes (on r) 1, 2 and 3 represent seasonal mean optical depth, seasonal total rainfall and seasonal mean wind respectively

the partial correlation eliminating the effect of wind ($r_{12.3}$) and rainfall ($r_{13.2}$) are also shown in the figure. It is interesting to note that the partial correlations with wind and rainfall are positive and negative respectively which are, in general, significant and are greater than the direct correlation. This is because, as indicated earlier, wind and rainfall themselves are correlated (the correlation coefficient is 0.54 which is significant). This result (as shown in Fig. 8) indicates the effectiveness of wind and rainfall as production and loss agents of aerosols at the coastal station, Trivandrum. Thus the seasonal variations in aerosol optical depths can be attributed atleast partly, to these mechanisms. As there are various other production and loss mechanisms for aerosols as discussed earlier, observations spread over a longer period would certainly help to delineate the effects of these sources/sinks in a more unambiguous manner.

Aerosol Columnar Size Distributions

Aerosol optical depth $\tau_p(\lambda)$ can be written as

$$\tau_p(\lambda)=\int^{r_2} \pi r^2 \, Q_{\mathrm{ext}}(r, m, \lambda) n_c(r) \mathrm{d}r,$$

where

r is radius of the aerosol particle (assumed spherical),
Q_{ext} is the Mie extinction efficiency factor,
m is the complex refractive index,
λ is the wavelengths at which τ_p is obtained,
and

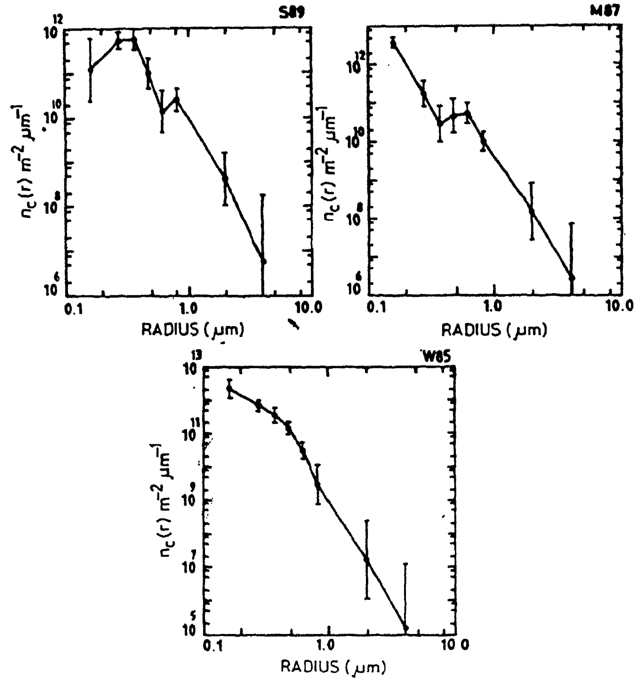


Fig 9 Aerosol columnar size distributions for summer 1989 (S 89), Monsoon 1987 (M 87) and winter 1985 (W85)

$n_c(r)$ is the total atmospheric aerosol content at radius r per unit radius interval.

By inverting the above equation, the columnar aerosol size distribution $n_c(r)\mathrm{d}r$ can be obtained. This equation is similar to the Fredholm integral equation of the first kind. Methods of inverting the above equation for obtaining $n_c(r)\mathrm{d}r$ are dealt with in literature[32,33]. Krishna Moorthy *et al.*[13] adopting the method of King *et al.*[32] derived the columnar aerosol size distributions from $\tau_p(\lambda)$ data at Trivandrum.

The columnar aerosol size distributions have been derived using the seasonal mean aerosol optical depth data at Trivandrum. In Fig. 9 are shown typical examples of the size distributions for the three seasons. The main features of the seasonal characteristics of the size distributions are the following:-

i) The size distributions are bimodal type in all the summer seasons (from November 1985 to February 1990) with the secondary mode at aerosol radius of ~ 1 μm.

ii) The winter size distributions are mainly unimodal or inverse power law type.

iii) The monsoon size distributions show variability in functional form in different years. They are either bimodal or unimodal or inverse power law type.

The appearance of a secondary mode consistently in summer is indicative of an additional source of aerosols. It is pointed out earlier that in summer

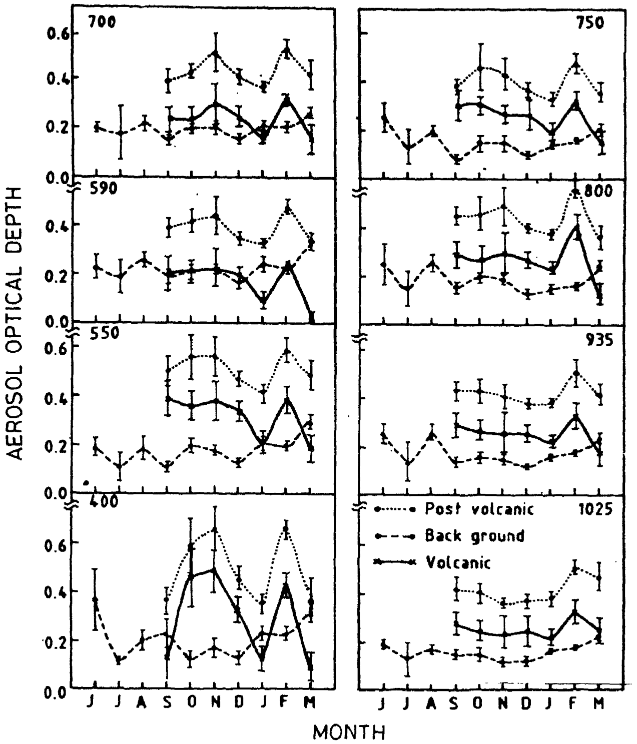


Fig 10 Temporal variations of monthly mean optical depths corresponding to post-volcanic, background and volcanic conditions

season, strong westerly wind is presently giving rise to sea surface agitation and consequent production of sea spray aerosols. The sea spray aerosols produced are known to be of size $>0.5\mu m$. The secondary mode appearing in summer season size distribution can be attributed mainly to this additional source of aerosols. In the monsoon, in addition to the production of sea spray aerosols, the loss process due to wet removal would be quite significant owing to heavy rainfall during this season. Thus, the nature of monsoon size distribution would depend upon the relative effectiveness of these strong production (sea spray aerosols) and loss (wet removal) processes. Thus, the observed variability in the monsoon size distribution for different years may be attributed to the variability in the relative strengths of these processes. The winter season is characterized by weak surface winds and insignificant rainfall. The photochemical processes leading to gas-to-particle conversion also become weak. Thus the winter aerosols can be considered to represent the background condition. When the main production and loss processes are weak, coagulation and gravitational sedimentation will be the dominant processes affecting the size distributions. Coagulation in the small particle range and gravitational sedimentation in the large particle range would result in a unimodal distribution which is observed, in general, in winter.

The total volume V_L of aersols is estimated by

$$V_L = \int^{r_2} r^3 n_c(r)\,\mathrm{d}r \qquad 8$$

It is found that partial correlation of V_L with wind eliminating the effect of rainfall is significantly positive and that with rainfall eliminating the effect of wind is significantly negative. This result further confirms the effect of wind and rainfall on aerosols discussed earlier in terms of aerosols optical depths.

Effect of Mt. Pinatubo Volcanic Eruption on Aerosol Optical Depths

Mt. Pinatubo Volcano (15°N, 120°E) situated in Philippines erupted in June, 1991, injecting large quantities of SO_2 gas and ash into the atmosphere. It was estimated that the ejected mass of SO_2 was about 20 million tones. This quantity of SO_2 converted to sulphuric acid aerosols in stratosphere, resulted in substantial (additional) aerosol loading.

Using the Multiwavelength radiometer at Trivandrum, the effect of the volcanic aerosol loading on aerosol optical depths has been investigated. Fig 10 shows the monthly mean aerosol optical depths at the different wavelengths due to post volcanic, background and volcanic aerosols for different months[4]. The background aerosol optical depths are obtained by averaging the aerosol optical depths for the same months in different years and the volcanic aerosol optical depths are obtained by subtracting the background aerosol optical depths from the post volcanic optical depths for different months.

It can be clearly seen that there is substantial enhancement in the aerosol optical depth following the volcanic eruption. The enhancement showed temporal variations with peaks in October-November 1991 and February 1992 with the strengths of these peaks varying with wavelengths. The size distribution of the volcanic aerosols have been estimated using this optical depth data. It is found that the size distributions generally exhibit a strong peak around 0.75μm and that the temporal changes in the size distribution occurred mainly in the small particle range. This can be attributed to the conversion of the small particles to larger sizes by coagulation. The volcanic aerosol mass loading (columnar) is found to be ~ 145 milli grams/m^2.

References

1 A Robock *Geofis Int* **23** (1984) 403
2 M C McCracken and F M Luther *Geofis Int* **23** (1984) 385
3 R K R Vupputuri and J P Blanchet *Geofis Int* **23** (1984) 433
4 K Labitzke, B Naujokat and M P McCormick *Geophys Res Lett* **10** (1983) 24
5 K Labitzke and M P McCormick *Geophys Res Lett* **19** (1992) 207
6 D J Hofmann *Rev Geophys* **25** (1987) 743
7 K Parameswaran, K O Rose, M Satyanarayana and B V Krishna Murthy *Indian J Radio & Space Phys* **18** (1989) 202
8 J D Klett *Appl Opt* **20** (1981) 211
9 F G Fernald *Appl Opt* **23** (1984) 652

10 K Parameswaran, K O Rose, B V Krishna Murthy, M T Osborn and L R Mcmaster *J geophys Res* **96** (1991) 10861

11 K Parameswaran, B V Krishna Murthy, K O Rose and M Satyanarayana *Ann Geophys* **11** (1993) 61

12 R P Turco, P Hamill, O B Toon, R C Whitten and C S Kiang *J atmos Sci* **36** (1979) 699

13 K Krishna Moorthy, P R Nair and B V Krishna Murthy *J appl Meteorol* **30** (1991) 844

14 P Hamill, O B Toon and C S Kiang *J atmos Sci* **34** (1977) 1104

15 G T S Bluth, S D Doiron, C C Schnetzler, A J Krueger and L S Walter *Geophys Res Lett* **19** (1992) 151

16 N D Sze and M K W Ko *Atmos Environ* **14** (1980) 1223

17 P J Crutzen *Geophys Res Lett* **3** (1976) 73

18 R P Turco, R C Whitten, O B Toon, J B Pollock and P Hamill *Nature* **283** (1980) 283

19 R P Turco, O B Toon, R C Whitten, E C Y Inn and P Hamill *J Geophys* **86** (1981) 1129

20 G K Yue and A Deepak *Appl Opt* **20** (1981) 3669

21 B V Krishna Murthy, K Parameswaran, K O Rose and M Satyanarayana *J atmos Terr Phys* **55** (1993) 809

R A Fisher *Statistical Methods for Research Works* Oliver and Boyd, Edinburgh (1970)

23 B V Krishna Murthy, K Parameswaran and K O Rose *J atmos Sci* **43** (1986) 914

24 R P Turco, R C Whitten and O B Toon *Rev Geophys Space Phys* **20** (1982) 233

25 B C Forster *Int J Remote Sensing* **5** (1984) 799

26 K Krishna Moorthy, P R Nair and B V Krishna Murthy *Indian J Radio Space Phys* **18** (1989) 194

Y P Rao Southwest Monsoon *Meteorol Monograph Synoptic Meteorology*, No 1/1976, India Meteorology Department, New Delhi (1976)

28 P K Das *Monsoons*, WMO-No 613 (1986)

29 V Narayanan *Indian J Meteorol Geophys* **18** (1967) 497

30 N S Nair and V Narayanan *Monsoon* **31** (1980) 409

31 C E Junge *Air Chemistry and Radioactivity* Academic Press, New York (1963)

32 M D King, D M Byrne, B M Herman and J A Reagan *J atmos Sci* **35** (1978) 2153

33 M D King *J atmos Sci* **39** (1982) 1356

34 B V Krishna Murthy, K Krishna Moorthy and P R Nair *Geophys Res Lett* (1993)

REMOTE SENSING OF CLOUD PRECIPITABLE WATER CONTENT USING INSAT-1D DATA

J K MISHRA AND O P SHARMA
Centre for Atmospheric Sciences, Indian Institute of Technology, New Delhi-110016, *India*
AND
M D JOSHI
Department of Civil Engineering, Indian Institute of Technology, New Delhi, India

(*Received* 21 *May* 1993; *Accepted* 13 *July* 1993)

Remote sensing of a cloudy atmosphere for the estimation of cloud precipitable water content (PWC) has been undertaken using INSAT-1D data during the monsoon month of July, 1991. The clouds have been classified using an unsupervised clustering technique based on the infrared ($11 \times 11 km^2$ resolution) and visible brightness ($2.75 \times 2.75 km^2$ resolution) characteristics. Satellite inferred cloud top temperatures have helped in estimating the top pressure and the geopotential height using the mean profiles for aerological parameters derived during the Monsoon Experiment (MONEX-79). The scheme estimates the cloud optical thickness, the liquid water path(LWP), the cloud base height and the PWC between the cloud levels. Satisfactory results for qualitative validation of base height estimation have been achieved. for quantitative validation, the estimated PWC has been correlated with the actual rainfall over numerous stations and a strong positive correlation (co-efficient of correlation equal to 0.8) has been found out. Two threshold limits viz., $PWC > 1.2 gm/cm^2$ for necessarily precipitating and $PWC < 0.35 gm/cm^2$ for never precipitating clouds, have been found out. The scheme gives accurate results for clouds which are large and really contiguous and can therefore be used efficiently for rainfall distribution analysis. Improvements are possible using data with better spatial and temporal resolutions and by modifying the scheme for night-time analysis as well.

Key Words: Cloud Classification; Cloud Base Height; Precipitable Water Content

Introduction

Conventionally, rainfall is estimated by analyzing the rainfall data collected using a dense network of rain-gauge stations. Daily rainfall is recorded and a central co-ordinating station is provided the data from all other stations. Isohyets are drawn and this information is used for hydrological studies. Lately, the meteorological satellites are being progressively used for both the estimation as well as the forecasting of rainfall. Remotely sensed imagery and digital data from the satellites offers an alternative method for this purpose. With the improvement in technology (in the spatial resolution of the radiometers), satellite data analysis has become quite a reliable, efficient and cost effective method.

Basically, there are two types of satellite observations which are used to estimate the precipitation (Arkin and Janowiak, 1991). These are: *i*) cloud top observations using infrared (IR) or visible brightness (VIS) or both (VIS/IR);

and *ii*) microwave observations of liquid and solid hydrometeors. These techniques have been labelled as indirect and direct respectively by Arkin and Meisner[2].

The indirect method utilizes VIS/IR data and aims to establish the correlation between VIS/IR characteristics and the actual rainfall. The cloud top temperature is usually determined using the infrared channel data. The associated cloud top pressure and the geopotential height can be determined using this top temperature. Cloud type identification still rests as a subjective procedure and no single classification is presently available. However, this cloud type information is useful in the initialization of the cloud water content for the estimation of cloud optical properties. The estimation of the base height using satellite data is of considerable importance. The ambiguities in the determination of cloud base height and the complications arising from the presence of precipitable water droplets are well known. This is because the ground observations usually provide reliable information for low clouds and for other clouds only in the absence of lower clouds. The high clouds, on the other hand, are easily sensed by the satellites. The middle clouds, therefore, have great uncertainties associated with them, especially when multilayered situations are present.

Several approaches relating the cloud optical properties with the satellite measured radiances[3-5] have been suggested. The liquid water path (LWP) is dependent upon the cloud top and base levels and the assumed liquid water concentration profile between these two levels. The optical thickness can be related to the LWP as given by the Stephens[3] parametrization scheme. The precipitable water content (PWC) can be estimated by integrating the mixing ratio profile between the cloud top and bottom pressure levels[6]. The correlation between the PWC and the actual rainfall can thereafter be established.

In this study, INSAT-1D (Indian Satellite) data from both the Infrared (IR) and visible (VIS) channels for 3 consecutive days in the monsoon month of July 1991 (23-25th), have been utilized. The suggested scheme identifies the cloud type, estimates the cloud top and base levels, the optical thickness, the LWP and finally the PWC between the cloud levels. Lastly, the correlation between this estimated PWC (over a scan spot/station) and the actual observed rainfall has been worked out.

Methodology

The cloud type classification is the first step if such a scheme is to be designed. Earth Resources Data Analysis (ERDAS) based image processing system has been used for this purpose. Infrared and visible data for all the three days have been classified using an unsupervised clustering technique. Next, using the scheme given by Lee and Taggart for visual cloud identification, numerous clusters have been combined into seven cloud types. The classification is based on the top temperature characteristics of each cloud. The spectral properties (infra-red and visible brightness) identified for each cloud type in this study is shown by a 2-dimensional diagram (Fig. 1).

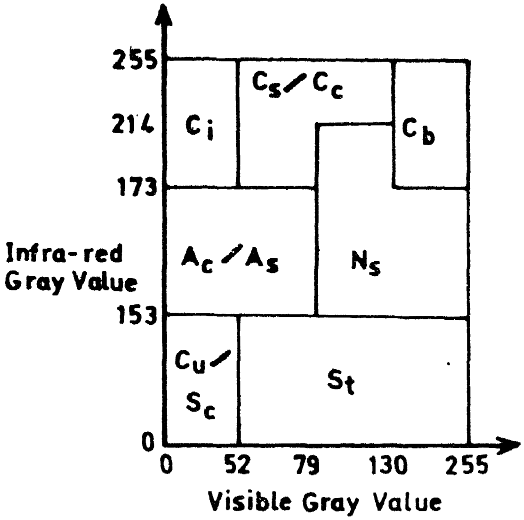


Fig 1 Cloud classification (using infrared and visible data) used in this study

Satellite infrared data are used for inferring the cloud top temperature by using a look-up table. This parameter is then used to estimate the top pressure and the geopotential height of each cloud. In our scheme, we assume that the cloud behaves like an air parcel which is in stable equilibrium with the surrounding atmosphere. This enables us to estimate the top pressure and the geopotential height using the mean profiles for aerological parameters in free atmosphere obtained during Monsoon Experiment (MONEX-79) as reported by Zaitseva[8].

Our next aim is to estimate the base height of the cloud. It may be pointed out here that till now we have fixed only the top height and have identified the cloud type for each scan spot. For the estimation of the base height we basically utilize the visible channel data for the optical thickness estimation which is then used to fix the base level of the cloud. The stepwise scheme is as follows:

Step 1: Estimate the optical thickness (τ) using the regression relation established between the visible data (VIS) and the optical thickness (explained later) given as:

$$\tau = 8.0625\left[\exp\left(\frac{0.0101\times VIS}{\cos\mu}\right)-1\right] \qquad \ldots (1)$$

Step 2: Estimate the LWP using the optical thickness (τ) using the relation given by Stephens[3]:

$$\log_{10}(LWP) = (0.5454\times\tau)^{0.254} \qquad \ldots (2)$$

Step 3: Estimate the base height (z_b) by solving analytically the relations employed in the National Center for Atmospheric Research (NCAR) model (Kiehl, 1990)[9]:

$$LWP=\int_{z_b}^{z_t} w\,\mathrm{d}z \qquad \ldots (3)$$

$$w=w_0 \exp\left(-\frac{z}{h}\right) \qquad \ldots (4)$$

$$h=300+2900\cos\phi \qquad \ldots (5)$$

It may be observed that an exponentially decaying profile (expression (3) for the liquid water concentration has been assumed in the NCAR model. The liquid water concentration (LWC) at a height z is given as w; z_t and z_b refer to the cloud top and bottom levels respectively. The maximum LWC is therefore at the base level and is given as w_0. The liquid water scale height (h), is a latitude (ϕ) dependent quantity. It is also pertinent to point out here that the cumulonimbus (Cb) clouds, which are more common during the monsoon period, usually have much higher LWC values (around 2.5×10^{-3}kg m^{-3} However, since the spatial resolution (11×11km^2) of INSAT-1D for the IR channel is quite large, the choice of such high LWC value could cause significant overestimation due to pixel averaging for the *Cb* clouds. To overcome this problem, the value of w_0 has been judiciously fixed at 0.2×10^{-3} kg m^{-3} for all the clouds in this study.

The regression relationship between the optical thickness and the visible reflectance, mentioned in Step 1 needs elaboration: the authors have derived the relationship between these two parameters by regressing the optical thickness limits of each cloud type (International Satellite Cloud Climatology Project (ISCCP)) radiometric definition of cloud type (Rossow, 1990)[10] with the visible gray level limits of the corresponding cloud type (classified by the image processing (IP) technique). We form pairs of points consisting of the optical thickness limits (ISCCP) and the corresponding VIS limits for each cloud type. For example, the optical thickness limits for *Cb* clouds are 23 (minimum) and 125 (maximum) and the corresponding visible gray level limits are 130 (minimum) and 253 (maximum). So, the pairs are (23, 130) and (125; 253). Similar pairs for the other cloud types are also formed. Next, all these pairs are used for obtaining an expression between the optical thickness and the visible gray level value by regression. The final relation (expression (1)) is given as:

$$\tau=8.0625\left[\exp\left(\frac{0.0101\times \mathrm{VIS}}{\cos\mu}\right)-1\right],$$

where μ is the solar zenith angle; VIS is the visible gray level value. The root mean square error (RMSE) for this expression is found out to be 1.02, which is very satisfactory for our purpose. However, note that the number of data pairs are rather small and therefore the RMSE expected is also small.

This stepwise procedure provides the likely base height which can be further used for PWC estimation. These base height are then validated with ground observations taken at the time coinciding with the satellite data (0600

UTC) However, there are possibilities of geometric errors in coordinate fixing and pixel averaging.

Next, the precipitable water content (PWC) is estimated. For this purpose we proceed as follows:-

The fractional precipitable water Δw contained between any two levels of the atmosphere, separated by a pressure interval of Δp, is given by:

$$\Delta w = \frac{r \times \Delta p \times 1000}{g} \qquad \ldots (6)$$

where r is the humidity mixing ratio in gm/gm of air and g is the acceleration due to gravity. We estimate this fractional precipitable water contained between the cloud top and bottom levels (which have been previously estimated) and refer to this quantity as PWC in the entire paper. The humidity mixing ratio is derived using the tephigram for near saturated condition. The PWC is therefore given as:

$$\text{PWC} = \sum_{\substack{\text{Cloud} \\ \text{Base } Pr}}^{\substack{\text{Cloud} \\ \text{Top } Pr}} r \times \Delta p \times \frac{1000}{g} \qquad \ldots (7)$$

In order to establish the correlation between this estimated PWC and the actual observed rainfall, we compare these two (PWC and rainfall) for 18 rain-gauge stations which are uniformly spread all over India. The 24-hour rainfall (observed) at 8:30 A.M. (0300 UTC) at each rain-gauge station is matched with the average PWC value (average of the estimated PWC over the station at 0600 UTC for the same day and the preceding day at 0600 UTC). Therefore, it has been assumed that the average PWC is representative for the entire period starting at the preceding morning till the next day morning (day on which 24-hour rainfall has been observed). The coefficient of correlation between these two has also been found out. The regions with fairly equal PWC distribution have been demarcated and these have been compared with the actual rainfall distribution.

Results and Discussion

The cloud type classification results derived using the present scheme have been validated with the ground observations. For this purpose the cloud types identified over a number of stations (15) using the VIS/IR classification (Fig. 1) have been compared with actual cloud type data (provided by Northern Hemispheric Analysis Center (NHAC), New Delhi) taken at the time coinciding with the satellite data. This comparison is shown in Table I. It may be observed that: *i*) in general the qualitative validation is reliable; *ii*) there is over-estimation of cloud top temperature (coldness) over atleast two stations., viz. Bombay and Bhopal; and *iii*) the usual problem of under-estimation of cirrus top temperature is observed over Madras station, as a result of which the clouds over Madras have been identified as middle level clouds (altocumulus) rather than high level cirrus clouds.

The estimated base height has been validated both qualitatively as well as quantitatively. Qualitatively, the base heights estimated using the scheme have been found to be in good agreement (Table I) with the usual classification of low, middle and high clouds (based on the base levels) viz. Low: 0-2km; Middle: 2-6km; High: > 6km. However, the base height of Stratus (St) clouds have been observed to be significantly higher than 2 km, thereby, requiring further investigations. The overall accuracy of the entire scheme i.e. a positive correlation of the PWC with the actual rainfall, indicates the suitability of the scheme for the estimation of base height itself and therefore, validates quantitatively.

The average PWC estimated over each scan spot (entire Indian region) for all the three days has been plotted. A comparison of the distribution of the estimated PWC and the observed rainfall distribution for all the days have been shown in Figs 2 (*a*) & (*b*), 3 (*a*) & (*b*) and 4 (*a*) & (*b*). It may be observed that the two corresponding figures are highly comparable.

In order to quantify these results, 18 stations (numbered in Figs 2, 3, 4) have been selected. Table II shows the comparison of the estimated PWC and the actual rainfall observed at these stations. The coefficient of correlation is found out to be + 0.79, which is quite significant. Threshold limits for precipitating and non-precipitating clouds have been identified and are as follows:-

PWC > 1.2 ------ always precipitating

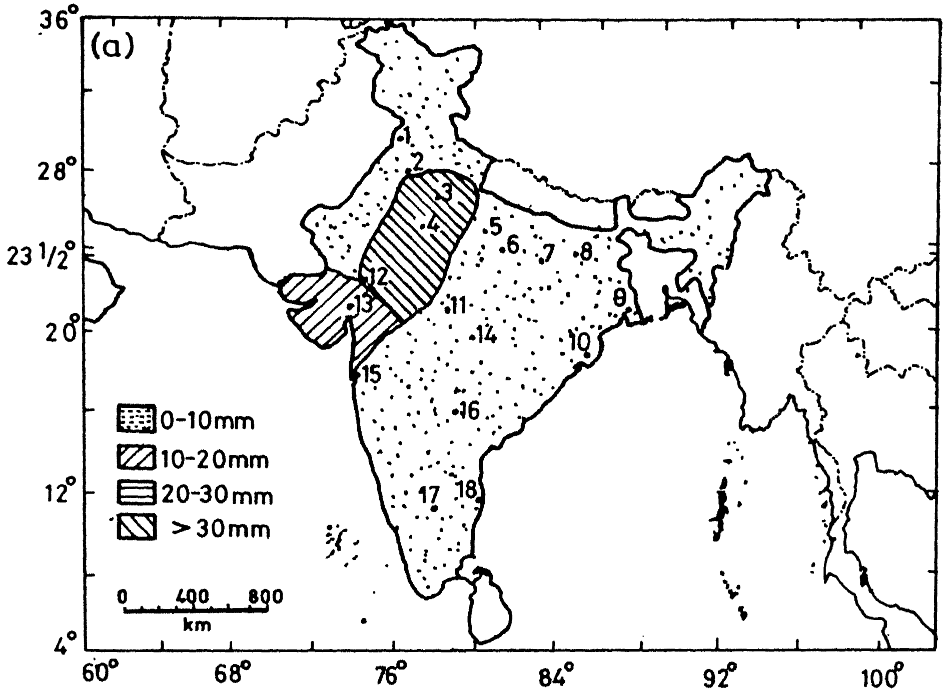
PWC < 0.35 ------ never precipitating

where all PWC values are in gm/cm^2.

Table I

Cloud type classification and cloud base height validations with ground observations

Station	Cloud Type (scheme derived)	Base Character (scheme derived)		Cloud Type (ground observation)
		Height (km)	Type	
Amritsar	C_S/C_U	7-08	High	A_C/C_I (High)
Delhi	A_C/A_S	4.98	Middle	A_C (Middle)
Jaipur	A_U/A_S	4.62	Middle	N_S/A_C (Middle)
Lucknow	C_U/S_C	1.71	Low	S_C (Low)
Allahabad	C_U/S_C	1.77	Low	S_C (Low)
Calcutta	C_S/C_U	8.08	High	$S_C/N_S/C_C$ (Mixed)
Bhubaneshwar	N_S	7.31	High	A_C/N_S (Middle)
Bhopal	C_B	3.62	Middle	S_T (Low)
Ahmedabad	C_B	4.13	Middle	C_B (Low)
Udaipur	C_B	4.02	Middle	C_B (Low)
Nagpur	A_C/A_S	4.11	Middle	S_T/N_S (Middle)
Bombay	N_S	4.8	Middle	S_T/C_S (Low + High)
Hyderabad	S_T	2.48	(Mid-low)	S_T/A_C (Mid-low)
Bangalore	S_T	2.56	(Mid-low)	S_T/A_C (Mid-low)
Madras	A_C	3.69	(Middle)	C_I (High)



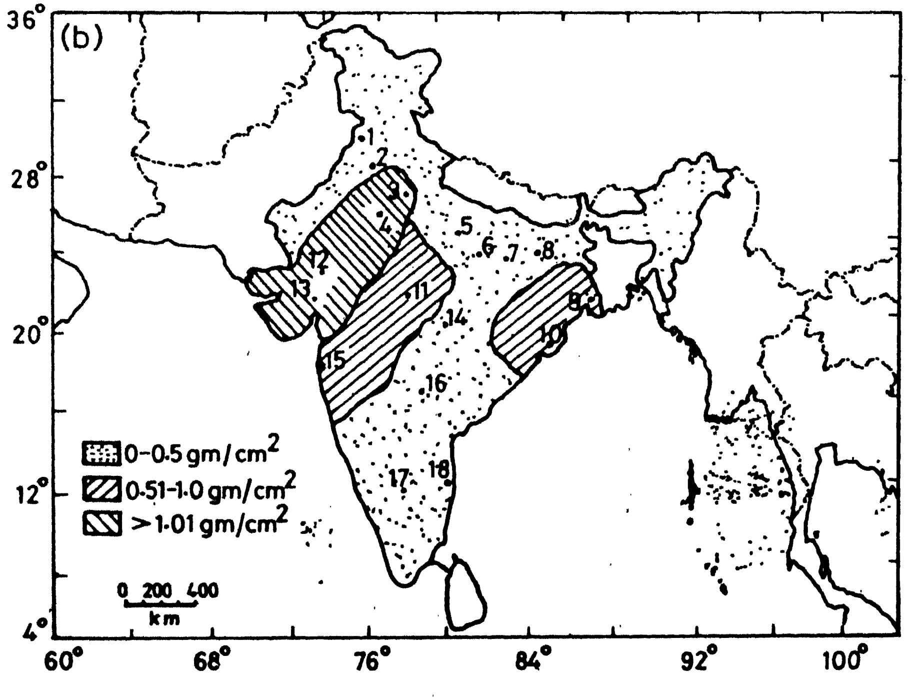
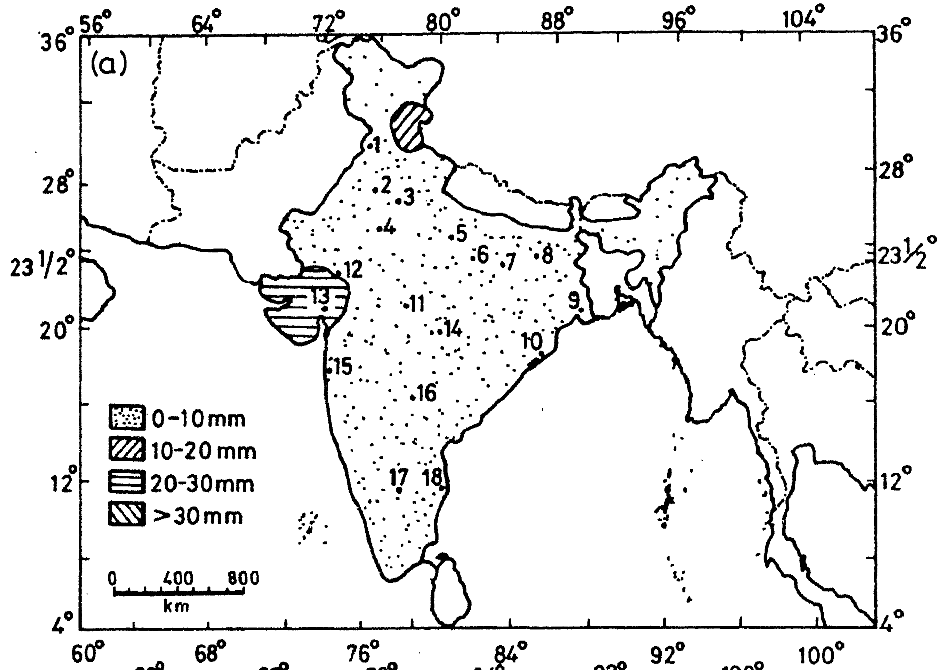


Fig 2(*a*) Rainfall distribution (actual) over India on 23-7-91

Fig 2(*b*) Precipitable water content (estimated) over India on 23-7-91



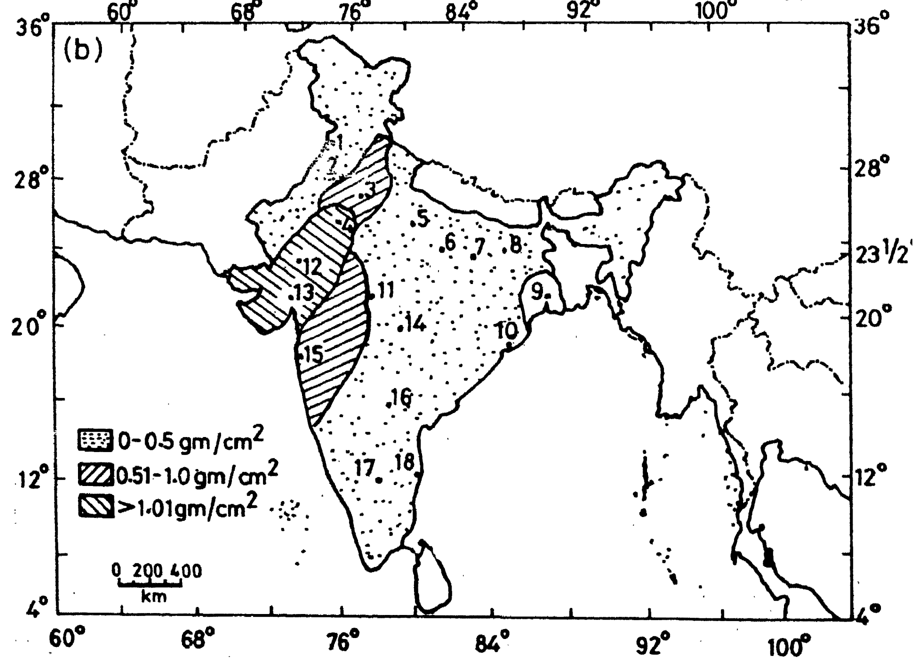
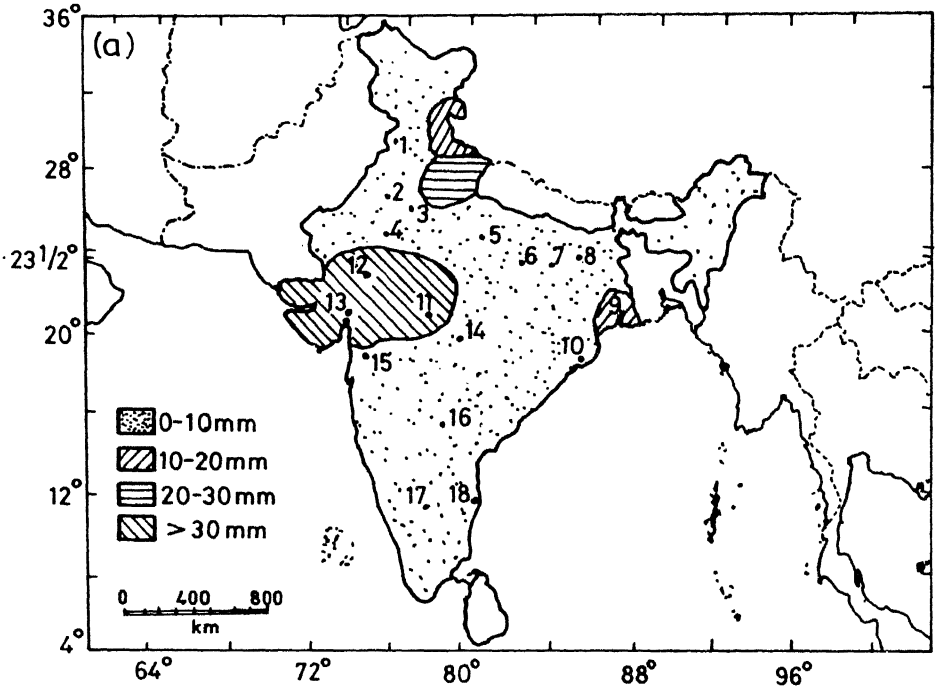


Fig 3(*a*) Rainfall distribution (actual) over India on 24-7-91

Fig 3(*b*) Precipitable water content (estimated) over India on 24-7-91



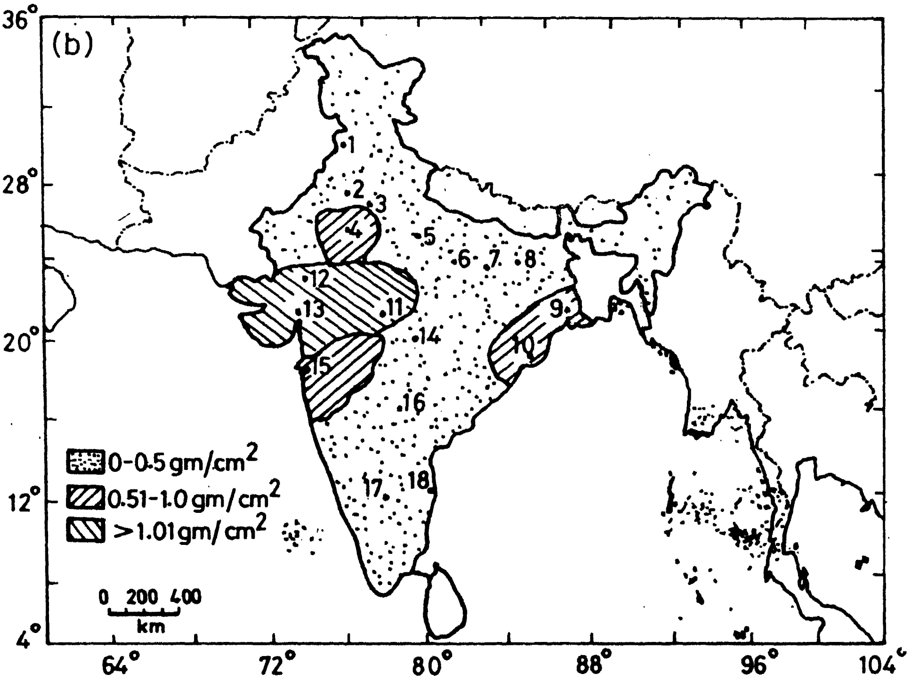


Fig 4(*a*) Rainfall distribution (actual) over India on 25-7-91

Fig 4(*b*) Precipitable water content (estimated) over India on 25-7-91

Table II

Comparison of estimated PWC and actual rainfall for all three days over the 18 stations

Station	23-7-91		24-7-91		25-7-91	
Name	PWC gm/cm^2	Rainfall mm	PWC gm/cm^2	Rainfall mm	PWC gm/cm^2	Rainfall mm
Amritsar	0.22	0	0.16	0	0.37	4
Hissar	0.30	0	0.23	0	0.54	0
Delhi	1.30	43	0.86	0	0.45	17
Jaipur	1.55	46	1.13	0	0.55	7
Lucknow	0.34	0	0.36	2	0.33	0
Allahabad	0.37	1	0.37	1	0.35	trace
Varanasi	0.30	trace	0.38	6	0.26	0
Patna	0.50	0	0.41	2	0.41	0
Calcutta	0.85	3	0.88	0	0.68	10
Bhubaneshwar	0.51	1	0.34	0	0.50	0
Bhopal	0.63	3	0.45	5	1.30	56
Udaipur	1.52	34	1.03	2	1.25	43
Ahmedabad	1.03	11	1.06	22	1.36	59
Nagpur	0.44	0	0.40	4	0.36	trace
Bombay	0.72	5	0.70	1	0.59	3
Hyderabad	0.37	0	0.37	0	0.33	trace
Bangalore	0.35	trace	0.35	trace	0.35	–
Madras	0.27	0	0.31	0	0.31	0

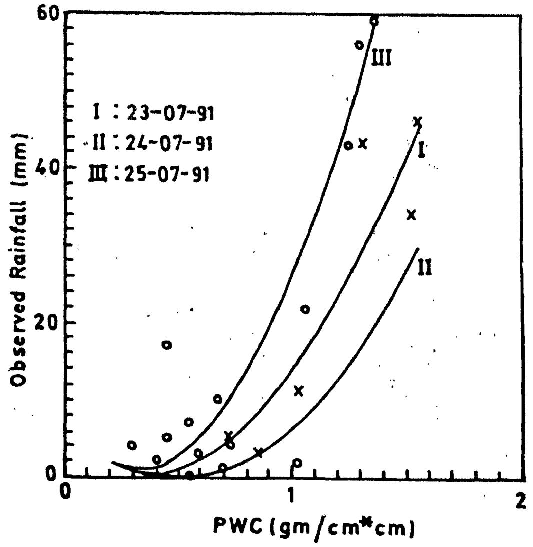


Fig 5 Actual rainfall versus precipitable water content estimated over 18 stations

This implies that whenever the PWC for any type of cloud exceeds 1.2 gm/cm^2 value it necessarily precipitates since it cannot hold any excess water. Also, the PWC has to attain a limit of at least 0.35 gm/cm^2, for any cloud type, so as to start precipitating.

Fig. 5 shows the curves for rainfall versus the estimated PWC for all the days. It can be observed that a curvilinear relationship (best fit curve is a polynomial of degree 2 exists between the two. It is also noteworthy that in the initial portion of the curves (till PWC < 0.5), there is little change in the rainfall. As an improvement to this analysis, we plot the curves for rainfall versus that value of PWC which is in excess of 0.5 gm/cm^2 (i.e., PWC-0.5) for only those values of PWC which are greater than 0.5 gm/cm^2. The corresponding curves are shown in Fig. 6. These curves show a better linear profile than the initial curves (Fig. 5). Fig. 7 shows the ellipse which is formed by enclosing all the plotted points of Fig. 6. The coefficient of correlation for this set is found to be 0.8, which is slightly better than the previous one (coefficient of correlation equal to 0.79).

Since the PWC has been found to be a good indicator of the actual rainfall, it can be used for actually estimating the rainfall. In that case, reliable and quick estimates of the rainfall distribution can be obtained in a cost effective manner. This scheme, using satellite data can therefore be conveniently used for this purpose.

In the scheme, errors are possible on a number of accounts: *i*) when clouds are smaller than the instrument's field-of-view (FOV); *ii*) geometric errors aris-

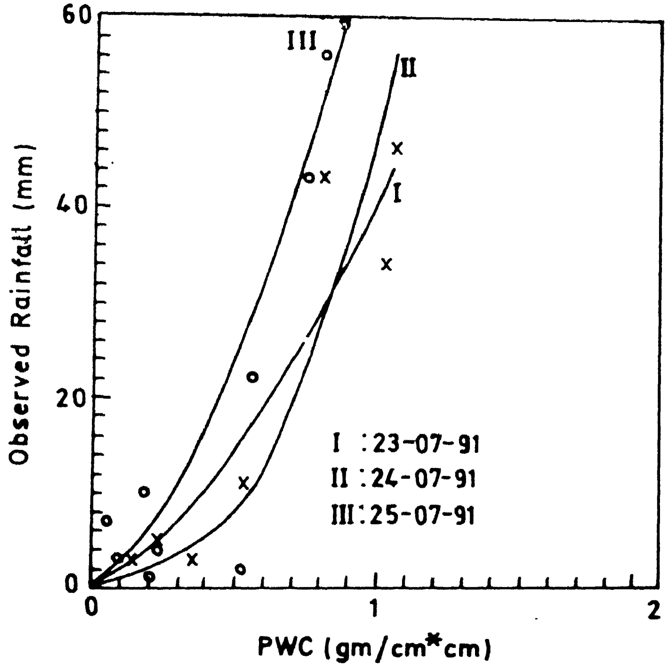


Fig 6 Actual rainfall versus (PWC-0.5) values for PWC values greater than 0.5gm/cm^2 only

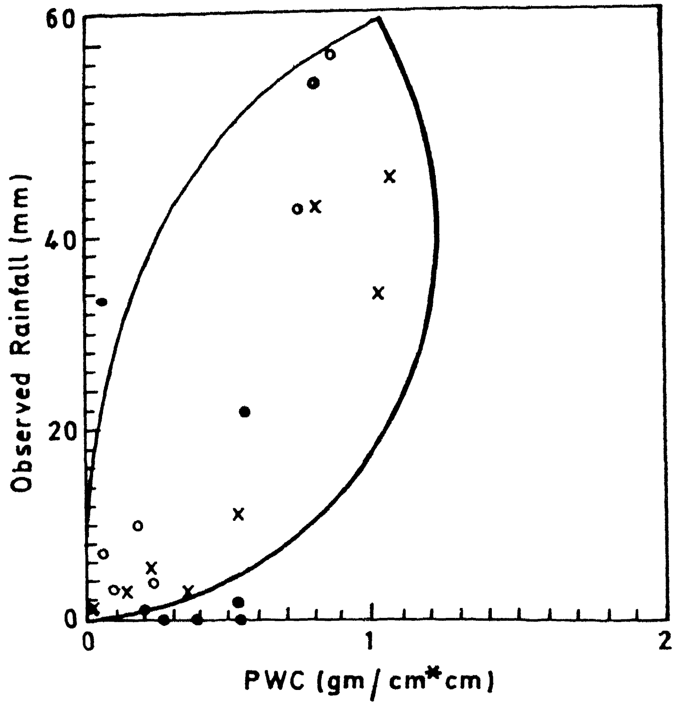


Fig 7 Ellipse formed by enclosing all plotted points of figure 6

ing due to exact co-ordinate fixing of the stations; and *iii*) errors in assuming the averaged PWC to be representative of 24-hour duration. The scheme can therefore be improved by utilizing 4-hour data for PWC estimation. Also, since VIS data is used for cloud type identification, this scheme needs to be modified for night-time analysis.

Conclusions

The cloud type classification scheme using unsupervised clustering technique yields fairly reliable results, especially when clouds are areally large and contiguous (as during the monsoon period). Errors in identification are observed when clouds smaller than instruments FOV are present. The under-estimation of cirrus top temperature is also observed. The base height estimation has been validated both qualitatively and quantitatively. The estimated PWC has been found to be strongly correlated (coefficient of correlation of 0.8) with the actual rainfall observed over 18 stations. There is a good resemblance between the PWC and rainfall distribution for entire India during the study period. Two threshold limits viz., PWC > 1.2gm/cm^2 for necessarily precipitating clouds and PWC < 0.35gm/cm^2 for non-precipitating clouds, have been found out. The overall accuracy of the scheme is satisfactory. It can, therefore, be used for reliable and quick estimates of the actual rainfall distribution. Further improvements are possible with finer spatial resolutions (1 × 1km^2) and temporal resolutions (4-hourly satellite data) and suitably modifying the scheme for night-time analysis.

Acknowledgements

We wish to thank the Director General of Meteorology and Dr R R Kelkar for making available to us all the data. We also express our deep sense of gratitude to Mr G S Mandal of N.H.A.C., New Delhi for his kind help. We are also indebted to Dr P Arkin of Climate Analysis Centre, U.S.A. for the discussions and his advice regarding this study.

References

1 P Arkin and J Janowiak *Analyses of the global distribution of precipitation Dyn Atmos Oceans* **16** (1991) 5-16
2 P Arkin and B N Meisner *Mon Wea Rev* **115** (1987) 2009-2032
3 G L Stephens *J atmos Sci* **35** (1978) 2123-2132
4 C M R Platt, D W Reynolds and N L Abshire *Mon Wea Rev* **108** (1980) 195-204
5 K T Kriebel, R W Saunders and G Gessel *Beitr Phys Atmos* **62** (1989) No. 3 165-171
6 S Rangarajan and A Mani *Proc Indian Acad Sci* (Earth Planet Sci) Vol 91 (1982) No. 3, 189-207
7 R Lee and C I Taggart In: *Climatology from Satellites* (E C Barret) Methuen & Co (1979) 90-92
8 N A Zaitseva, R M Vilfrand and M A Bathan *Results of Summer Monex Field Phase Research (Part-B) FGGE Operations Rep Ser* **9** (1980) 56-61
9 J T Kiehl *ECMWF/WCRP Workshop on Clouds, Radiative Transfer and the Hydrological Cycle* (1991) 49-59
10 W B Rossow *ECMWF/WCRP Workshop on Clouds, Radiative Transfer and the Hydrological Cycle* (1991) 49-59

ROLE OF TIME DELAY IN CLIMATE RESEARCH

KAUSTUBHA BHATTACHARYA

NCMRWF, MAUSAM Bhavan Complex,
Post Box No. 3103, *Lodi Raod, New Delhi-*110 003, *India*

(*Received* 14 *October* 1993; *Accepted* 15 *October* 1993)

The two crucial roles played by time delay in theoretical climate research are discussed. One is the well-known use of delays in computations of attractor dimension by embedding methods. The other is the ability of delay parameter in a delay differential equation to bring about bifurcations when they are appropriately changed. The latter case is illustrated by results from an one dimensional energy balance model with delayed albedo feedback.

Key Words: Delay; Embedding; Bifurcations

Introduction

The concept of time delay is very important in non-linear dynamics in general and climate research in particular. Ruelle[1] and Packard *et al.*[2] first showed how to use delay coordinates in building up of an embedding space for estimation of attractor dimension. With a scalar time series $X_n(t)$ sampled $N+1$ times at interval τ, i.e.

$$t = 0, \tau, 2\tau, ..., N\tau, \qquad \ldots (1)$$

the delay vector

$$X_n = (X(n\tau), X((n+1)\tau), ..., X((n+N)\tau)) \qquad \ldots (2)$$

can easily be constructed. Provided that the dimension (m) of such an embedding space is large enough (i.e., $m > 2d + 1$, where d is the dimension of the attractor) the attractor set of a dynamical system can be safely embedded into it (Whitney[3]; Takens[4]; Fraedrich[5]). Recently, this concept has been utilized by Grassberger and Procaccia[6] very effectively to design an algorithm for attractor dimension estimation. Although we will present this procedure in some detail, we will first discuss the delays in the natural climate system and the sources thereof.

There are several components of the terrestrial climate system. The most important of these are the atmosphere, the hydrosphere, the cryosphere and the biosphere. Of these the behaviour of first three can be quantified and their characteristic time scales are known to be very different, i.e., these components are delayed with respect to one another. It was Lorenz[7] who first pointed out the possibility of non-linear interaction between these different time scales resulting in irregular fluctuations that resemble the time series of the terrestrial climate. Such climatic fluctuations have been referred to as the climatic change due to internal causes, or more simply, as the internal climatic change.

There are many physical sources of delay in the terrestrial climate system, some of which can be easily identified. Ice, for instance, can change on widely different time scales, starting from snow (~seasons) to continent sized ice sheets (~millions of years). Similarly, oceans exhibit different time scales of change, with the upper oceans changing on seasonal time scale and the deep oceans changing on the time scale of thousands of years. Bare land has the fastest time scale of response. It has even been postulated there is a phase lag in the development of cumulus clouds with interesting consequences (Davies[8]).

This paper describes an one dimensional energy balance climate model with delayed albedo feedback. It is shown how this model responds to changes on different time scales.

Delays in Attractor Dimension Calculation

The basic concept of a delay vector is quite old and had been developed by statisticians. However, Whitney's[3] embedding theorem is new in that it gives some quantitative theresholds so that attractor reconstruction using the delay coordinates is a valid procedure. This reconstruction has, however, turned out to be hard to realize in practice, because measured time series are very often noisy and of short duration (Vautard *et al.*[9]). So much so that the very first task of deciding whether a given time series is indeed chaotic or just noisy is often a hard one (Provenzale *et al.*[10]). When both are present the problem is of course magnified many times. Another outcome of the noise is that a great deal of effort has been spent towards noise reduction. One method of reducing noise is through appropriate choice of basis vectors, where the vectors are determined by the time series itself. Usually the basis is chosen to be the eigenvectors of two-point correlation matrix calculated from data. This method, originally called Principal Component Analysis by the statisticians, is given different names: Singular Value Decomposition, Singular Spectrum Analysis Proper Orthogonal Vectors. The eigenvectors themselves are often called Karhunen-Loeve vectors or Empirical Orthogonal Functions.

Another important area of uncertainty is determination of optimal m or τ. Although Whitney's[3] theorem guarantees validity of attractor reconstruction as long as $m > 2d+1$, the attractor dimension d is not known in advance. Several methods exist for optimal computation of embedding dimension (m) or the delay (τ).

Although in principle any delay τ can be used, in practice the delay cannot be too small because in that case the delay vectors are hardly distinguishable from one another and it also cannot be too large because the delay vectors become statistically independent. A good (and usual) choice is to plot the autocorrelation as a function of τ, and find first point at which the autocorrelation is zero. Although this method suffices as a "thumb rule", it is intellectually more satisfying if such a criterion can be obtained from information theory which leads to a nonlinear notion of independence (Abarbanel *et al.*[11]). Fraser and Swinney[12] suggested that we compute average mutual information $I(t)$ defined as

$$I(t) = \sum P(X(n), X(n+t)) \log_2 \left[\frac{P(X(n), X(n+t))}{P(X(n)\, P(X(n+t))} \right] \qquad \ldots . (3)$$

where

$P(a, b)$ is the joint probability of measurement of a and b within their respective ensembles.

Then the first minimum $I(t)$ is a good choice for lag time τ. However, like the linear case this is also only a prescription.

Attractor Reconstruction

Starting from the original time series $X_n(t)$ we construct (Ruelle[1]; Packard *et al.*[2]) $N+1$ new time series by applying to each term of the series shifts of

$0, \tau, 2\tau, \ldots\ldots, (N-1)\tau, N\tau$

Grassberger and Procaccia[6] gave the following widely used algorithm: compute values of the (linear) correlation function $C(r)$ defined as

$$C(r) = \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} H(r - |\mathbf{X}_i - \mathbf{X}_j| \qquad \ldots (4)$$

where $H(x)$ is the Heaviside function.

A log-log plot of $C(r)$ against r yields a straight line. Starting from a small values of n these straight lines are successively plotted, and the slopes reach a plateau after an initial increase. This limiting slope is the attractor dimension (d), and the value of n for which this limiting slope is attained gives the minimum number of variables needed to fully describe the system.

One Dimensional Energy Balance Climate Model: Role of Delay

A delay differential equation is infinite dimensional in nature, and it is expected that a wide range of behaviours will be exhibited by such an equation. This was confirmed by an one dimensional energy balance climate model of the Sellers[13] type. This model was extended by Bhattacharya[14] *et al.* using some realistic assumptions that are valid over ice-age time scales. Here, the model is described briefly, details may be found in Bhattacharya *et al.*[14].

In this model latitude is the only space variable, and the energy balance at the earth's surface is described by the time-dependent equation

$$G(X) \frac{\partial T}{\partial t}(X, t) = R_i[X, T(X, t)] - R_0[T(X, t)] + F \left[X, t, \frac{\partial T}{\partial X}(X, t), \frac{\partial^2 T}{\partial X^2}(X, t) \right] \qquad \ldots (5)$$

where $X=(2\Phi)/\pi$, and Φ is the colatitude. $T(X,t)$ is the surface temperature and $G(X)$ is the heat capacity of earth's surface. On the R.H.S the successive energy terms are the absorbed part of the incoming solar radiation at the surface, the outgoing longwave radiation at the surface, and the meridional heat transport. They are represented as

$$R_i\, \mu\, Q(X)\{1-\alpha[X, T(X,t)]\} \quad \ldots (6)$$

$$R_0 = C[T(X,t)]\cdot \sigma \cdot [T(X,t)]^4 \quad \ldots (7)$$

$$F=\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}[\sin\phi\cdot K(\phi)]\frac{\partial T}{\partial\phi} \quad \ldots (8)$$

Two new features were added to the basic Sellers'[13] model, and they affect the results greatly. Hence they are discussed here in some detail.

The first is the inclusion of delay in the albedo formulation. As indicated earlier, this is the single most important reason for the interesting behaviour of the model to be described later. Whenever, snow or ice is present at the surface the albedo is determined not only the by the current temperature but also by weighted past temperatures. This weight is a maximum at $t=\tau$, and has a cutoff at $t=2\tau$. In absence of any guidance from the past records, the weighing function has been taken to be of Gaussian shape, with the peak of the curve at the current time.

The second is the addition of a kink in the temperature-albedo curve. This is located at that temperature where baroclinic activity normally takes place, and this is thought to coincide with the edge of the large icecaps. This hypothesis is partially supported by observations of cloudiness at the icecaps (Lamb[15], Schwerdtfeger and Kachelhoffer[16]). However, although qualitatively plausible, this relation is again difficult to formulate; so in an ad hoc manner a triangular shaped jump was used.

Two other minor features were a new computation of global surface heat capacity and restriction of heat transport to a linear formulation that is determined by the present temperatures only.

Results

Time independent (steady-state) version of the model yielded five solutions for the present climate, as against three for the original Sellers[13] model. We well denote them by $T_1(X)$, ..., $T_5(X)$. $T_1(X)$, $T_3(X)$ and $T_5(X)$ are linearly stable, while $T_2(X)$ and $T_4(X)$ are linearly unstable. Moreover, $T_1(X)$ represents the present (interglacial) climate, $T_3(X)$ represents the glacial climate, and $T_5(X)$ represents the "deep freeze"—a hypothetical climate where the earth was entirely covered with ice. All these have been elaborated upon in Bhattacharya *et al.*[14].

Results obtained from the time-dependent version of this model emphasize the dominant role played by the delay parameter in a delay differential equation. For all the integrations performed the starting temperature and albedo were consistent with each other. From start to $t=2\tau$ the time lag tail was built

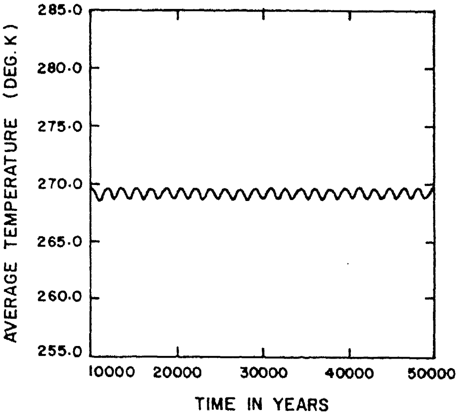


Fig. 1a

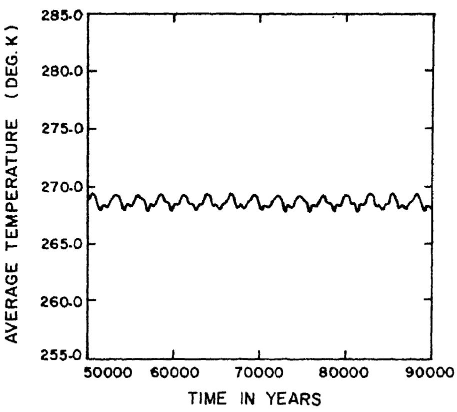


Fig. 1b

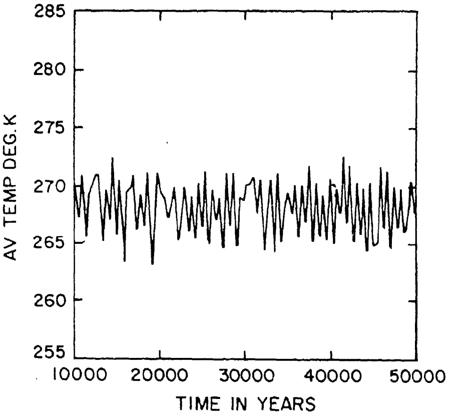


Fig. 1c

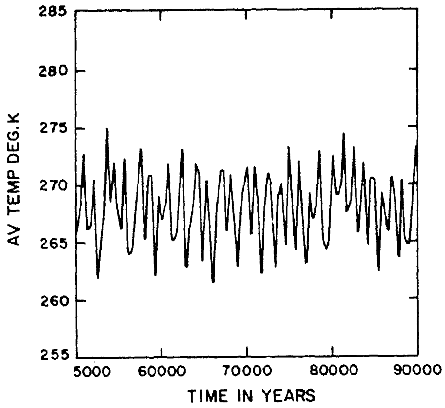


Fig. 1d

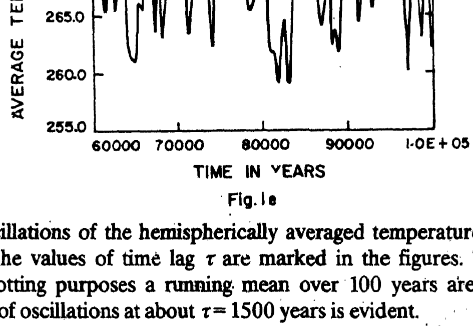


Fig. 1e

Figs 1*a*-*e* Sustained oscillations of the hemispherically averaged temperature $\overline{T}(t)$ around climate 1 are shown. The values of time lag τ are marked in the figures. Time units are in 1000 years. For plotting purposes a running mean over 100 years are shown. The change in the character of oscillations at about $\tau = 1500$ years is evident.

up by keeping the albedo fixed at the starting value. Thereafter, the albedo was allowed to be determined by past temperatures as described before. No oscillations were obtained when the starting temperature were equal to or higher than for present climate.

Sustained oscillations were obtained for negative perturbations that were large but not too large, i.e.,

$$T_1(X) - 10\mathrm{K} < T(X,0) < T_1(X) - 7.5\mathrm{K} \tag{9}$$

In these cases the shape and amplitude of the oscillations were determined by τ and σ, where σ is the half width of the Gaussian distribution that determines the past albedo from the past times. The actual oscillations for $\tau = 500$, 750, 1000, 1500, and 5000 years are shown in Figs 1*a*-*e* where the hemispherically averaged temperature $\bar{T}(t)$ are plotted against the time t. It is seen that between $\tau = 1000$ and $\tau = 1500$ years a qualitative change in the nature of the oscillations take place from periodic to aperiodic. For $\tau = 500$ years, the oscillations are simply periodic, while for $\tau = 750$ and 1000 years the oscillations are quasiperiodic. Clearly the change in values of the time-lag parameter τ has brought about the bifurcations of the system from simply periodic to quasi periodic, and then to chaotic.

Discussions

We have seen two important roles of time delay in climate research. One pertains to the usefulness of time delay in dimension calculations of the attractors, the main motivation here being the detection of strange attractors which is a non-trivial task for a short and noisy time series. The other is the role of the delay parameter in a delay differential equation. We have seen that by properly changing τ we can bring about bifurcations in such an equation which is an infinite-dimensional system. However the exact manner in which this happens is yet to be explained and that is subject of future research.

References

1 D Ruelle *Publications Mathematiques of the Institut des Hautes Etudes Scientifiques* **50** (1979) 27
2 N H Packard, J P Crutchfield, J D Farmer and R S Shaw *Phys Rev Lett* **45** (1980) 712
3 H Whitney *Ann Math* **37** (1936) 645
4 F Takens *Lecture Notes in Mathematics* **898** (1981) 366
5 K Fraedrich *J atmos Sci* **43** (1986) 419
6 P Grassberger and I Procaccia *Phys Rev Lett* **50** (1983) 346
7 E N Lorenz *J appl Met* **9** (1970) 325
8 H C Davies *Q J Royal met Soc* **105** (1979) 325
9 R Vautard, P Yiou and M Ghil *Physica D* **58** (1992) 95
10 A Provenzale, L A Smith, R Vio and G Murante *Physica D* **58** (1992) 31
11 H D I Abarbanel, R Brown, J J Sidorowich and L S Tsimring *Phys Rev A* (*Preprint*)
12 A M Fraser and H L Swinney *Phys Rev A* **33** (1986) 1134
13 W D Sellers *J appl Met* **8** (1969) 392
14 K Bhattachara, M Ghil and I L Vulis *J atmos Sci* **39** (1982) 1747
15 H H Lamb *Q J Met Soc* **81** (1955) 172
16 W Schwerdtfeger and S J Kachelhoffer *Antarctic J US* **8** (1973) 234

ATMOSPHERIC CLIMATE MODELLING—A VIEW OF CURRENT ACTIVITIES AT MPI HAMBURG*

L BENGTSSON AND L DUMENIL
Max-Planck-Institut für Meteorologie, Bundesstrabe 55,20146 *Hamburg, F.R.G.*

(*Received* 18 *January* 1994; *Accepted* 31 *January* 1994)

Introduction

The third scientific group at the Max-Planck-Institute for Meteorology was established on 1 January. 1991. It was successively built up during 1991 and in addition to the Director it consists of 6 scientists and 6 support staff. The group works closely together with the Climate Modelling Support Group set up by the German Climate Computing Centre (DKRZ) and the group of Klaus Hasselmann at MPI which focuses on the development and validation of coupled ocean/atmosphere models.

The methodology of the research is mainly directed towards the use of large scale comprehensive climate models and a considerable part of the work of the group is devoted to the development of an advanced atmosphere general circulation model (AGCM) including sub-models for land-surface processes and biomes.

The MPI meteorological climate modelling group has the following main scientific objectives:

- To further the physical and dynamical understanding of natural and forced fluctuations in the climate system with emphasis on the timescale 10^{-1}-10^{2} years. This research has the overall aim to obtain a more in depth knowledge of the effect of increasing greenhouse gases on the climate system.
- To explore the predictability of coupled ocean-atmosphere systems.
- To undertake climate simulation studies, with the objective to identify the regional climate effects (Europe, North Atlantic region) due to the increase in the atmospheric concentration of greenhouse gases and other external forcing of the climate system.

In a joint effort, work is under way to evaluate the capability of El-Nino type predictions by a coupled ocean/atmosphere model. Work is also in progress to incorporate chemical processes with emphasis on the sulphur, methane- and ozone cycle.

Atmospheric Model Development

A major effort by the group has been the development of a global general circulation model, ECHAM3, which has been used for most of the numerical experimentation work during the period. ECHAM3 is based on the ECMWF 19-level spectral-transform model but has in several respects a modified physical parameterization, in particular concerning the handling of cloud- and radiation processes. Cloud water, for example, is handled as a separate 3-dimensional variable. A detailed description can be found in Roeckner *et al.*[1] and an up-to-date model documentation is available from DKRZ.

---

*Belongs to Section I. Late receipt, hence added at the end of the volume.

ECHAM3 has been developed and programmed in a flexible way and has been used in long-term integrations in T21, T42, T63 and T106 horizontal resolution. It has also been used in multi-year ocean-atmosphere coupled experiments.

Parallel to the experimentation with ECHAM3, a new AGCM, ECHAM4, has been developed and test integrations are in progress. ECHAM4 has a new radiation parameterization including the effect of the "exotic" greenhouse gases, CH_4, NO_2, CFC etc. Boundary layer processes are handled by a higher order turbulence closure parameterization and the advection of moisture and liquid water are treated by a semi-Lagrangian integration. An arbitrary number of tracers can also be handled by the semi-Lagrangian scheme (see 'Global Transport Modelling' on p. 436). The model can be extended to at least the stratopause using a uniform radiation code.

Experiments have been undertaken with a full semi-Lagrangian integration, but so far mass-conservation has been unsatisfactory when integrated over longer time-periods. A new long term experiment is in progress using a T63 resolution; this integration is expected to be available within a month or so.

For the study of regional processes, a high resolution, limited are grid point model, originally developed for short-range weather prediction in the Nordic countries, has been modified for climate research. This has been achieved by inserting the physical parameterization of ECHAM3 and increasing the number of vertical levels to be the same as in ECHAM3. The new model, HIRHAM, has been coupled to ECHAM and numerical experiments are in progress. HIRHAM is also programmed in a flexible way and can be used for arbitrary areas of the earth and for resolutions down to 10-15 km. HIRHAM will also be used in parameterization studies. Areas under consideration are Indonesia and Antarctica.

Work is also in progress to couple a non-hydrostatic model, GESIMA, developed at GKSS, Geesthacht, to HIRHAM and thus to obtain a series of atmospheric models from global scale down to the kilometerscale.

The model development work takes place in close co-operation with other German groups as well as groups in Europe and United States.

Modelling Research

Dynamical Regionalization

General circulation models are the main tools available today for climate change simulations. However, for the very long integrations of 100 years or so normally required in climate applications, models can only be run at relatively low horizontal resolutions. Although a doubling of the resolution used so far may be possible within the next few years, it is still too coarse to adequately describe the regional climate details, required for climate change impact studies. Higher resolution simulations for time slices of the order of 10 years are possible and have been undertaken. Still higher resolutions down to 50km or less, however, can be reached by nesting a high resolution limited area model (LAM), covering a particular area of interest, into a course resolution global model. To pursue the latter possibility, a LAM, suitable for climate simulations has been developed in close cooperation with DMI and KNMI, the Danish and Dutch meteorological Services. This LAM, called HIRHAM, was developed by combining the physical parameterization package from the climate model ECHAM with the Nordic, Dutch, Irish operational forecasting model, which is called HIRLAM (High Resolution Limited Area Model). In close cooperation with DMI, the HIRHAM will be used in a series of experiments aiming at a verification of the nesting LAM-GCM technique for high resolution present day climate simulations over Europe. It is planned to test also a coupled atmospheric-oceanic LAM obtained by coupling HIRHAM with a limited area version of the OPYC ocean model.

These experiments which are part of an international project called "Regionalization" supported by the CEC and coördinated by B. Machenhauer, were started at MPI in January 1993. Several other projects in which HIRHAM might be used are presently under consideration. The increased resolution obtained with HIRHAM using a grid length of about 50km compared with the driving T42 ECHAM3 is illustrated in Figs. 1 and 2. The figures show for each model its orography and a 6 hour accumulated precipitation field. The corresponding synoptic situation as simulated by HIRHAM is illustrated by the surface map Fig. 3.

Design of a New Semi-Lagrangian Scheme

For short and medium range weather forecasting a shift from Eulerian to semi-Lagrangian integration techniques has taken place at several forecasting centres. The main reason for this development has been the increased computational efficiency of semi-Lagrangian models, but also an increased accuracy in the calculation of advection terms.

Due to problems with mass conservation in present semi-Langrangian models, climate modellers have so far hesitated to use these models, although increased resolution could be obtained for the same cost. Artificial corrections can be made each time step which prevents a systematic loss of mass during long integrations. Such corrections are, however, unsatisfactory, at least theoretically, as the mass which is added each timestep is evenly distributed over the globe and not added at the right geographical positions. Furthermore other important conservation properties of the continuous meteorological equations are not formally fulfilled in the present semi-Lagrangian formulations. How serious this is for long climate simulations is not known at present.

In order to assure an exact fulfilment of the most important integral constraints in the basic equations a new semi-Lagrangian formulation has been worked out. The idea is to use as prognostic equations Lagrangian forms of the continuity equation, the thermodynamic equation, the total energy equation and the angular momentum equation. These equations are discretized for vo-

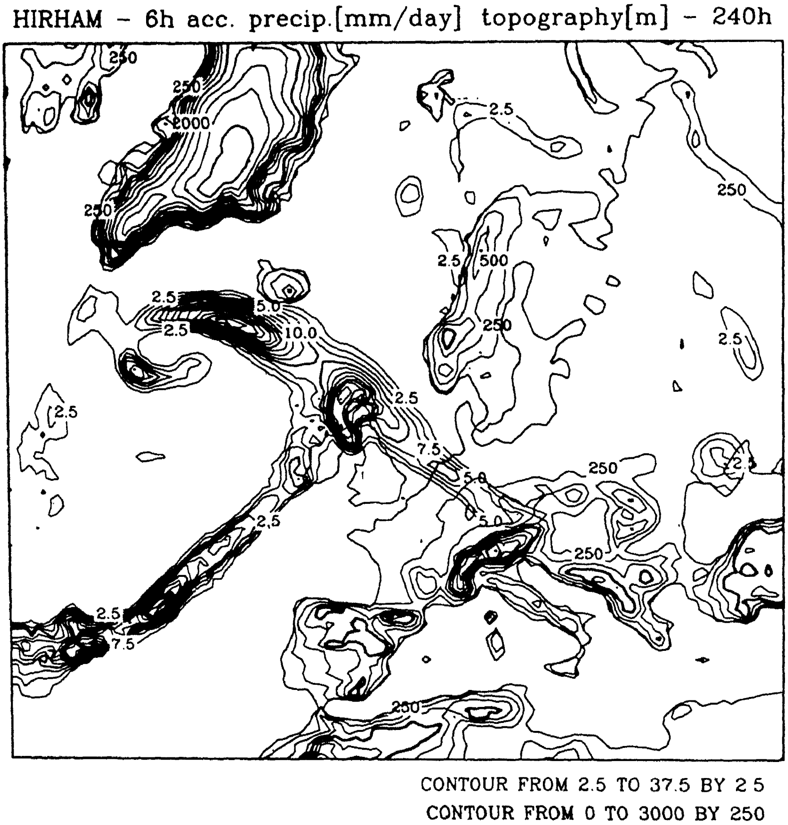

HIRHAM - 6h acc. precip.[mm/day] topography[m] - 240h
250
2000
2.5
5.0
10.0
7.5
500
CONTOUR FROM 2.5 TO 37.5 BY 2.5
CONTOUR FROM 0 TO 3000 BY 250

Fig 1 HIRHAM topography and 6 h accumulated precipitation. Contour lines for orography, every 250 m, and for precipitaion, every 2.5 mm.

lumes of air that are being advected in the horizontal wind field at time t and are ending up after each time step (at time $t+\Delta t$) in the grid volumes. The continuity equation is used to define the top and bottom pressure of each air volume at $t\Delta t$ and t. The explicit leap frog time extrapolation scheme chosen must be absolutely stable since the total mass, entropy, energy and angular momentum are conserved (except for external processes) over each time step. This formulation will be tested at first for simplified geometries (1-D and 2-D) using the shallow water equations.

To test whether present semi-Langrangian formulations with mass corrections can be used for climate simulations, a comparison of three pairs of 5 years runs, each with an Eulerian and a semi-Langrangian model, is presently being done. These runs will be made with the ECMWF, NCAR and the Canadian (RPN) models.

Stratospheric Modelling

The aim is to construct a comprehensive model of the general circulation of the troposphere and stratosphere with a uniform physical parameterization. The starting point is the ECHAM4 general circulation model which uses semi-

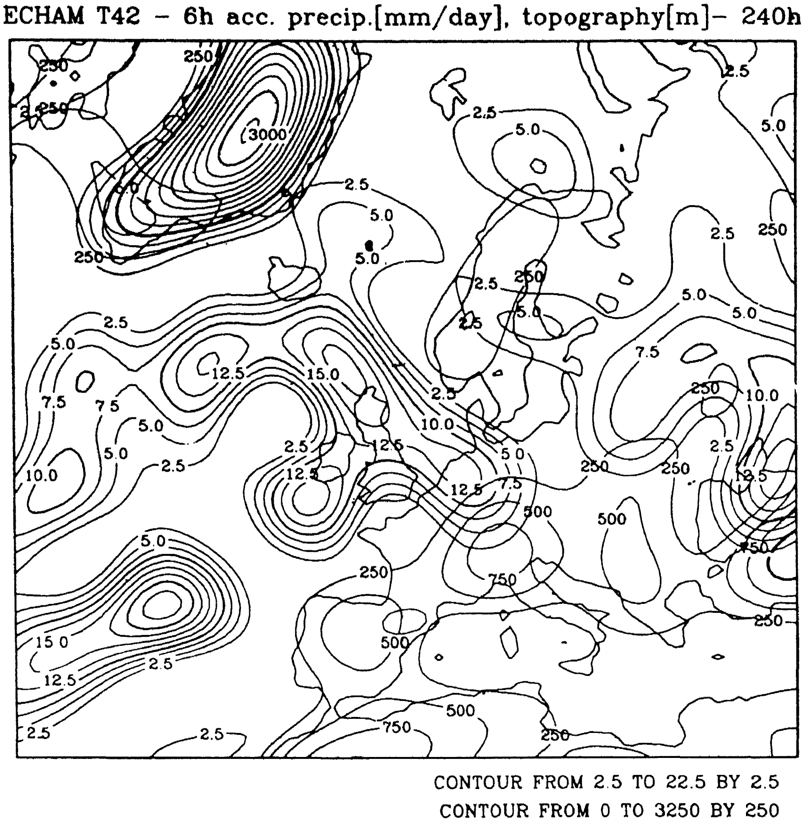

ECHAM T42 - 6h acc. precip.[mm/day], topography[m]- 240h
CONTOUR FROM 2.5 TO 22.5 BY 2.5
CONTOUR FROM 0 TO 3250 BY 250

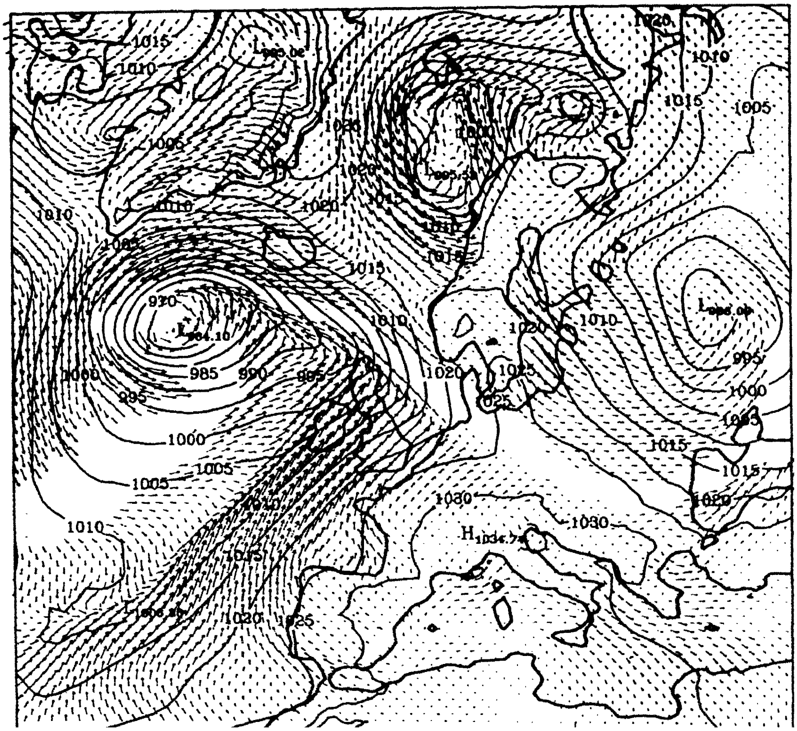
Fig 2 The same as Fig 1 but for ECHAM3 at T42 horizontal resolution.

Langrangian transport of water vapour and liquid water and a new radiation scheme.

The original ECHAM4 model has 19 vertical levels (hybrid vertical coordinate) with the top full-level pressure at 10 hPa (about 32km). A new vertical discretization was therefore necessary. In order to incorporate the whole stratosphere for such purpose an algorithm that generates vertical structures of the vertical coordinate has been constructed. This new algorithm allows a smooth increase of the spacing between adjacent vertical levels. Among the various vertical discretizations possible the following two have been selected: L25, with 25 vertical levels and top full-level pressure at 0.3 hPa (about 57km); and L35; with 35 vertical levels and top full-level pressure at 0.1 hPa (about 65km). In the upper troposphere and lower stratosphere L25 has the same vertical resolution of L19, while L35 has finer resolution than L19.

In constructing the new comprehensive model the physical parameterization has been uniformly extended through the stratosphere. However, it has been found that the approximate method to compute the longwave radiative heating rates in between full radiation time steps used in ECHAM3 fails in the upper stratosphere. This problem has been solved keeping the net longwave radiative flux constant between full radiation time steps. Since the standard full

HIRHAM – 10m wind, 850hPa temperature – 240h



CONTOUR FROM 965 TO 1035 BY 5

Fig 3 HIRHAM simulated wind at 10 m and temperatures at 850 mb. Colour changes for every 4°C.

radiation time step is two hours only, it appears acceptable to neglect longwave flux variations over this time-interval. Without any other basic modification of the parameterization scheme, three integrations of the ECHAM4 model with L19, L25 and L35 levels respectively have been performed. Each integration has been carried out for 360 days with annual mean forcing (i.e. constant solar radiation, prescribed sea surface temperature and ozone distribution) and T21 spectral truncation. As initial state an isothermal, motionless and dry atmosphere has been used. Although some improvement is noted in the representation of the lower stratosphere (weakening and euqatorward shift of the stratospheric westerly jets), the results show the necessity of a revision of the radiation scheme above 10 hPa. Such a modification has now been done including the incorporation of the Voigt band correction.

Boundary Layer Investigations

Research on *scale aggregation over type-A landscapes*, i.e., terrain with changes in surface conditions at scales *smaller* than 10km, has focused on extending the so-called concept of blending height to include form drag due to roughness elements. Based on recent work, which indicates that small areas of large roughness or minor obstacles dominate the regional momentum flux, whereas the regional heat and moisture fluxes are determined by the dominant surface cover, it has been hypothesized that regional momentum flux should be evaluated from an effective roughness length which implicitly accounts for the form drag of obstacles, whereas regional heat and moisture fluxes should be estimated from local surface parameters such as local roughness lengths. This hypothesis compares favourably with other, but less flexible, proposals when testing it in numerical simulations (Claussen 1992)[2].

The same hypothesis is applied when computing regional surface fluxes over partly forested areas. Here, neglect of air flow into and from the edges of tall vegetation appears to result in underestimation of local advection. On a regional scale, this 'edge effect' leads to an increase of momentum flux, but a decrease of latent heat flux (Claussen and Klaassen, 1992)[3].

The concept of blending height has been used to estimate atmospheric forcing in a large-scale sea-ice-oceanic mixed-layer model around Antarctica. The atmospheric drag on sea ice is computed from an effective roughness length which implicitly takes into account skin friction on ice and water as well as form drag on ice floe edges. Furthermore, the atmospheric drag depends on regional thermal stability. It turns out that, particularly in the Wedell and Ross Seas, an upward heat flux is diagnosed, although the atmospheric surface layer is stably stratified at the regional scale—an illustrative example of the so-called Schmidt's paradox. The new parameterization leads to improved model results in terms of ice-thickness distribution, ice velocities, mixed-layer depth and mean oceanic heat flux (Stössel and Claussen, 1992)[4].

Research on *scale aggregation over type-B landscapes*, i.e. terrain with variation in surface characteristics at scales *larger* than 10km, is in progress. It seems that some ideas valid for scale aggregation at small scales can simply be extended to larger scales. Although further testing has to be done, a simple version has been tested in ECHAM3. First results show that horizontal gradients of near-surface temperatures at the east coasts of continents during winter are smoothed leading to less extreme changes in heat fluxes. Morevoer, the monthly mean isotherms of near-surface temperatures look more realistic than in the original model, where, in many cases, they simply follow the coarse land-sea mask.

Cloud-Radiation Processes and Their Implication on Climate

In ECHAM, stratiform clouds are simulated on the basis of a cloud water transport equation including sources and sinks due to condensation, evaporation and precipitation formation by coalescence of cloud droplets and sedimentation of ice crystals. The predicted cloud water content is used to parameterize the cloud optical properties such as emissivity and optical depth.

Satellite data of the Earth Radiation Budget Experiment (ERBE) are used to validate the model simulations, in particular the cloud radiative forcing (CRF) which is the contribution of clouds to the shortwave cooling and longwave warming, obtained by subtracting the respective clear-sky estimate from the total. According to Fig. 4-a-c, there is generally good agreement between

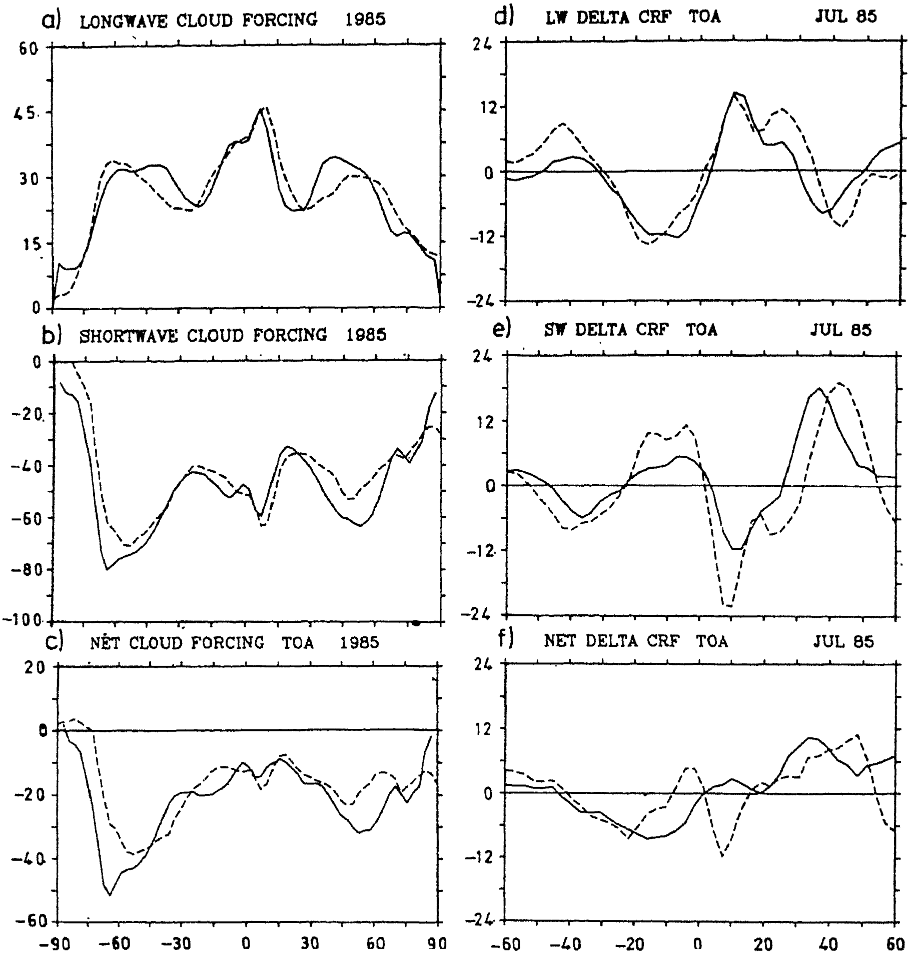


Fig 4a-f Annual zonal mean cloud forcing (Wm^{-2}) as analysed from ERBE data (full lines) and as simulated by ECHAM3-T42 (dashed). (a-c)
Response of the cloud forcing (Wm^{-2}) to seasonal thermal forcing for July 1985, as analysed from ERBE data (full) and as simulated by ECHAM3-T42 (dashed). Positive values correspond to a warming relative to the annual mean and negative values correspond to a cooling. (d-f)

the simulated (ECHAM3-T42) and observed CRF for 1985 (the model is forced with the respective SST for that particular year), and the deviations are mostly within the range of observational error which is around 10 W/m^2 for zonal means. On the regional scale (not shown) the deviations are larger and more significant. For example, the shortwave CRF is overestimated in areas of persistent tropical convection and underestimated over the northern oceans during the summer season due to a lack of low-level cloudiness.

Satellite data can also be used to validate cloud variations resulting from observable forcing such as the seasonal or ENSO cycles. As an example, Fig. 4d-f compares the observed and simulated CRF response to seasonal thermal forcing (July 1985 vs. annual mean 1985). Both, model simulation and ERBE data indicate a net mid-latitudinal warming of the summer hemisphere with respect to the annual mean (i.e., an amplification of the seasonal warming by the change of cloudiness), caused by a reduced shortwave CRF as a result of the seasonal transition of cloudiness over the Northern Hemisphere continents from a stratiform regime in winter to a convective regime in summer. In the tropics, the change of the cloud forcing (both components) is just a response to the seasonal shift of the ITCZ. While the model is able to reproduce the observed CRF response with some success, there is clearly an overestimation of the shortwave response in the tropics. Accordingly, the sign of the net response is reversed (damping of the seasonal wave instead of a slight amplification as suggested by the ERBE data). However, these preliminary results should be taken with caution because interannual variability and observational errors have to be considered as well.

It is well known that the atmospheric circulation in the tropics is largely driven by the release of latent heat in deep convective towers. On the other hand, very little is known about the radiative impact of these clouds on both the atmospheric circulation and the oceanic heating. In a collaborative effort together with the SCRIPPS Institution of Oceanography (Professor V. Ramanathan) a series of model experiments has been designed in order to study the interaction between cloud-radiative effects and the tropical circulation. Preliminary results obtained from this ongoing work (which will gain observational support from the CEPEX experiment in the tropical Pacific during March 1993 as well as TOGA-COARE) may be summarized as follows:

- In areas of persistent convection, as for example over the warm pool in the West Pacific, the net radiative heating by clouds (i.e. the atmospheric contribution to the cloud radiative forcing) contributes 20-30% to the total diabatic heating.
- Most of the radiative heating is provided by the trapping of upwelling infrared radiation by optically thick cirrus anvils above 400 hPa.
- Neglecting the radiative heating by the cirrus anvils causes a substantial weakening of the tropical circulation (Hadley and Walker cells), see Fig. 5.
- The most sensitive area in this respect is the warm pool area in the West Pacific.

According to Houghton *et al.* (1990), the current estimates of the equilibrium temperature response to a doubling of CO_2 range between approximately 3 and 5°C in the global annual mean. Most of these uncertainties are attributed to different climate sensitivities of the models, caused primarily by large differences in the simulated cloud feedbacks. A recent model intercomparison study initiated by Professor R.D. Cess aimed at comparing the $2 \times CO_2$ forcing rather than the climate feed-backs and suggests that a sizeable fraction of the simulated temperature range (up to 50%) may be explained by the different radiative forcing. The CO_2 forcing problem has been addressed in a joint effort among 14 GCM groups (including MPI). For computational reasons the models are run in a perpetual July mode. The radiative transfer code is called 4 times during a time step (pair of $1 \times CO_2$ and $2 \times CO_2$, with and without clouds respectively). In this simulation 3 of the 4 heating rate calculations are

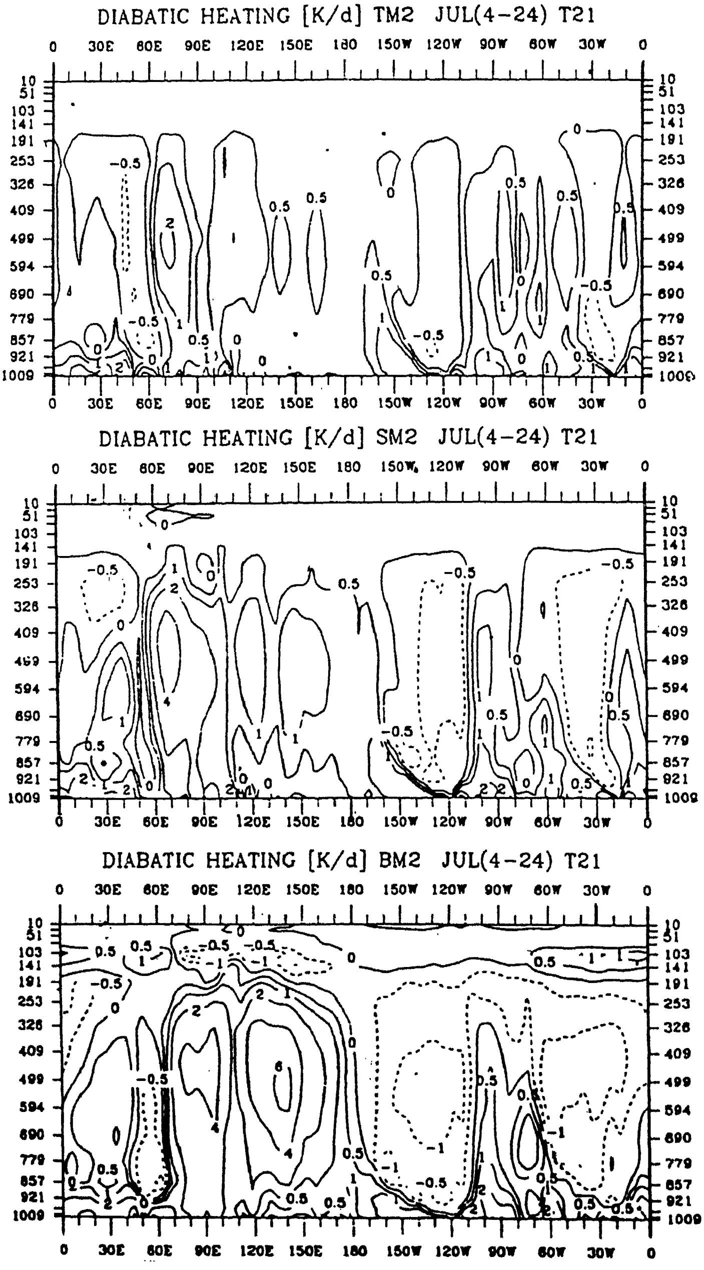
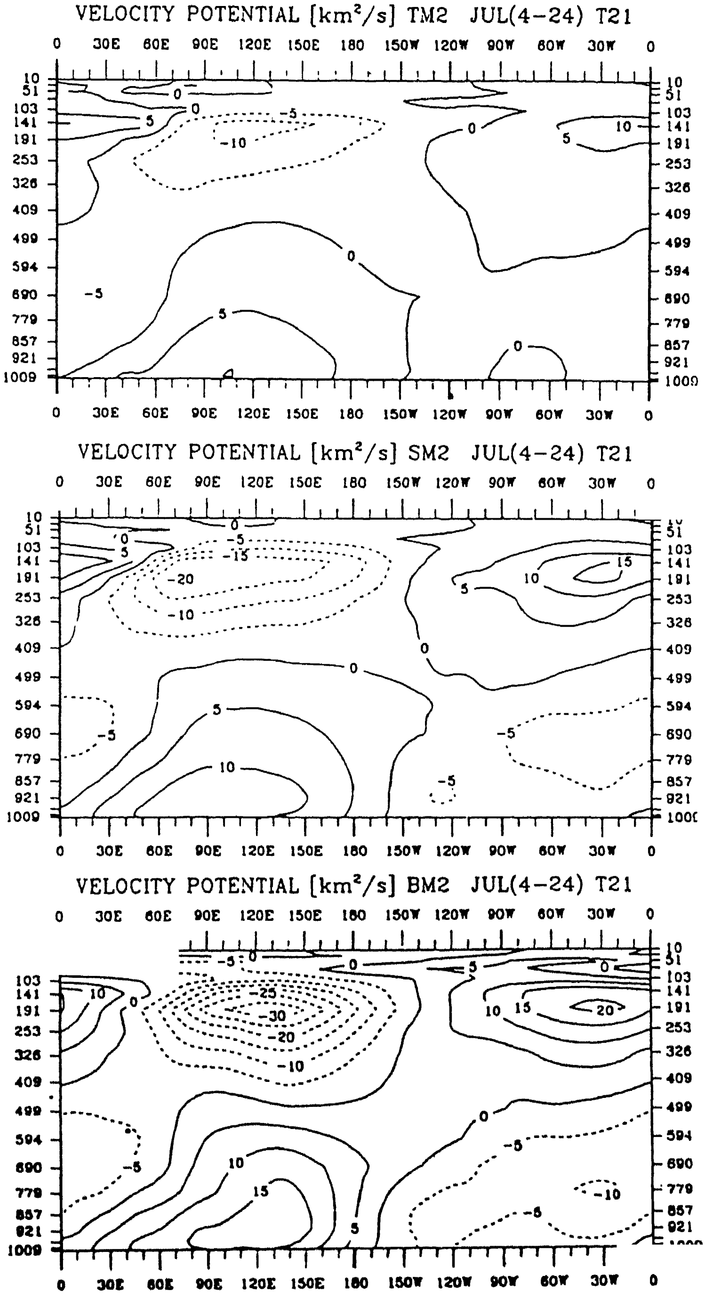


Fig 5 Longitude (0-30°N)/height (pressure).

Cross sections of total diabatic heating and velocity potential as obtained from three 24-month perpetual July simulations with the ECHAM3-T21 model with different assumptions on cirrus cloud emissivity:

(Continued on page 417)



(Fig.5 contd.)

Top: Transparent cirrus in the infrared

Middle: Variable cirrus emissivity depending on cloud water content (Control experiment)

Bottom: Black cirrus

just performed to diagnose the forcing without changing the model climate. Additional column experiments are performed for mid-latitude summer conditions and the results are compared with exact line-by-line integrations. The main conclusions from this intercomparison study are:

- The range of uncertainty is about 30% of the mean tropospheric forcing which is close to 4 w/m^2. Most of this uncertainty comes from the different treatment of the CO_2 absorption bands in the various models. Overlap effects due to water vapour and clouds and solar absorption do not contribute tot he variations in the forcing estimates.
- As compared to line-by-line models, the broad-band radiative transfer codes used in the 14 GCMs tend to underestimate the longwave $2\times CO_2$ forcing (mid-latitude summer, clear sky). Part of this underestimation is due to the neglect of certain CO_2 absorption bands in the broad-band models.

Development of Rainfall Run-off Scheme

In order to allow a meaningful estimate of the continental runoff which is used as fresh water input into the ocean in coupled ocean/atmosphere integrations, a rainfall/runoff scheme (Dümenil and Todini, 1992)[5] and a river routing scheme (Sausen, Schubert and Dümenil, 1992)[6] were developed. The rainfall/runoff component is a simplification of a state-of-the-art approach which is employed in hydrological modelling. The scheme takes into account the heterogeneity of the land surface within a grid area. It is the first application of such a model in a GCM. It is combined with a global data set on soil water holding capacities. The model is being intercompared together with more sophisticated hydrological models currently being used in other GCMs in the framework of the EC project on Spatial Variability.

The Global Runoff Data Centre (GRDC) has been approached with a request to provide data from their archives for model validation. At a first instance, longtime series of measurements for the 10 largest rivers of the world was selected. This has now been extended to include the 50 largest rivers (Dümenil *et al.*, 1993)[7]. Time series and long-term averages of monthly means are now available. The data provide an independent and valuable data set for validation of the hydrological cycle in the model. Model verification for individual basins is in progress.

Validation Studies

Results from several 10-year integrations with and without observed sea surface temperature, SST, have been investigated with respect to systematic deviations of the model runs from reality. These investigations cover a wide range of studies including the validation of:

(i) the large scale dynamic circulation;
(ii) the interannual variability;
(iii) regional climates;
(iv) physical processes including certain aspects of the hydrological cycle; and

(v) phenomenological evaluations such as blocking, extra-tropical cyclone tracks, tropical cyclones etc.

Some result of this work is reported in Roeckner *et al.* (1992)[1]. Several other aspects of the validation of ECHAM3 is presently being written up and hence only a summary of this work will be reported here.

Large Scale Systematic Errors

The monthly averaged or seasonally averaged circulation agrees well with observations except in a few areas. For the Nothern Hemisphere winter the following could be mentioned:

(i) The Atlantic subtropical anticyclone has a more easterly position and the Azore high has a centre over Spain. In connection to this the east Mediterranean trough is too weak.

(ii) The Aleutian low is positioned too far eastwards and is also too confined in the north-south direction.

(iii) The trough over the western Pacific and the ridge over Alaska at 500mb is too weak.

However in making these remarks, it should be noted that at several winters such features are actually observed in nature the only difference is that they appear more often in the model calculations.

The high as well as the low frequency variability is realistic, Fig. 6, although there exist a slight tendency to underestimate low frequency variability and overestimate high frequency variability.

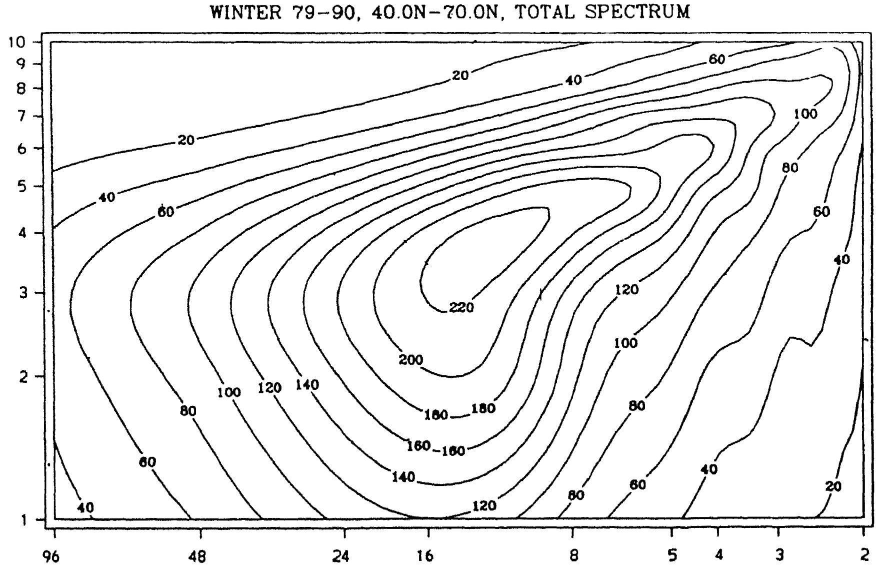
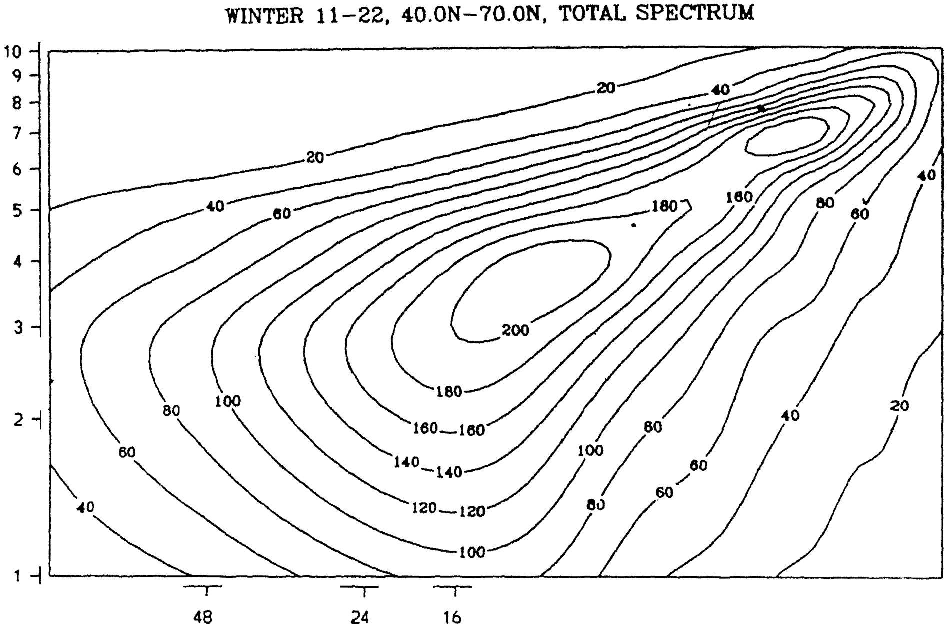


Fig 6(a) Power spectrum of geopotential at 500 hPa and 40° N-70° N for winter derived from 12 years of ECMWF analysis. Units are 100 m^2, on the abscisa the period (days) and on the ordinate the zonal wavenumber are given both in logarithmic scale. The spectral estimates were multiplied by wavenumber and frequency. Winter was defined as a period of 96 days starting December 1.



WINTER 11-22, 40.0N-70.0N, TOTAL SPECTRUM

(b) The same as 6a but for 12 years of an ECHAM3 experiment using observed SST.

The interannual variability at middle and high latitudes agrees well with ECMWF's analyses the effect of observed SST's in the Tropics describes more than 50% of the variance, but hardly any effect at all can be found in extra-tropics. See further 'SST-Impart Studies' that follows later.

The overall climatology of the simulations has been studied in its regional details. This has been accomplished by expressing the results of the simulations in terms of Köppens classification (Lohmann *et al.* 1992)[8] and also to calculate what biomes the model simulates. (*vide* also section on 'MONEG-Experimentation').

Such validation do efficiently highlight specific model deficiencies of particular importance for climate impact studies examples are too high precipitation over Australia and South Africa during the Southern Hemisphere summer. Over the Northern Hemisphere there is a tendency to simulate too little precipitation over land during the summer and in some areas, like the central part of United States the surface temperature in summer is also too high. As is being discussed below it is not clear to what extent this is caused by the handling of land-surface processes or whether the problems are related to deficiencies in the large scale circulation such as dynamical forcing caused by a too high precipitation over central America (in all likelihood due to orographic,representation difficulties at T42 resolution).

Low Frequency Variability

In order to distinguish processes influencing the atmospheric variability by spatial and temporal scales a space-time spectral analysis was performed. As a result the spectral distribution of both stationary and transient eddies was given. It is found that the "ultralong" transient waves (zonal wavenumber 1 to 3) give major contributions in the time range longer than 10 days, whereas those with a shorter wavelength (zonal wavenumber 4 to 7) have their maximum at periods shorter than 10 days.

As a consequence the data were subjected to a low- and a band-pass filter, the first one retaining eddies on time scales beyond 10 days, the latter one between 2 and 10 days. Thus we could distinguish the part of the atmospheric variability resulting from low-frequency and higher frequency disturbances. Another band-pass filter retaining eddies on time scales between 2 and 6 days, commonly used in many studies, was also applied to the data. Both band-pass filters lead to similar spatial structures, the 2-to-10 day filter, however, had significantly higher variance.

The distinction between the low- and higher frequency eddies as described above was also considered in the localized "Eliassen-Palm Flux" diagnostics. Thus we could investigate the impact of the transient eddies on the zonal mean flow. Those, for example, are thought to play an important role in initiating and maintaining a "blocked" flow and other quasi-stationary features.

As a prominent example of a low-frequency phenomenon blocking was investigated. Based on different theoretical descriptions various criteria were applied in order to identify episodes, when a blocked flow occurs in the extratropics region. It seemed that the criterion based on the "persistence" of such a phenomenon was the most appropriate one for our purposes, in particular as it can be applied on both hemispheres, but can also be used to identify other low-frequency phenomena in the extratropics such as "Cut-off Lows".

As a result individual episodes of Blocking were defined for particular locations leading to frequency distributions of this phenomenon distinguishing by seasons. Some of these episodes, however, will also be subjected to more detailed diagnostics considering important physical processes leading to a blocking flow regime.

As was seen from the spectral analysis the ultralong westward travelling waves give a major contribution to the atmospheric variability in the time range beyond 10 days. For that reason a first attempt was made to investigate a possible relation between the activity of these waves and the occurrence of blocking.

The Hydrological Cycle in the ECHAM3 Model

The group is spending considerable efforts in validating the hydrological cycle. Comparison with different climatological estimates show good agreement and in fact in many respect the different climate estimates differ between themselves as much as to the models results. Fig. 7 shows the estimated and simulated hydrological cycle for the Earth. Fig. 8 shows the same for the zonally averaged precipitation summer and winter respectively. Intercomparisons have also been undertaken against satellite estimates from OLR and SSM/I measurements as well as against land station data compiled and analysed by GPCC (Global Precipitation Climatological Centre) in Offenbach. GPCC has made data available for 1987 and 1988. These data have been compared with model integrations for the same period (using observed SST data).

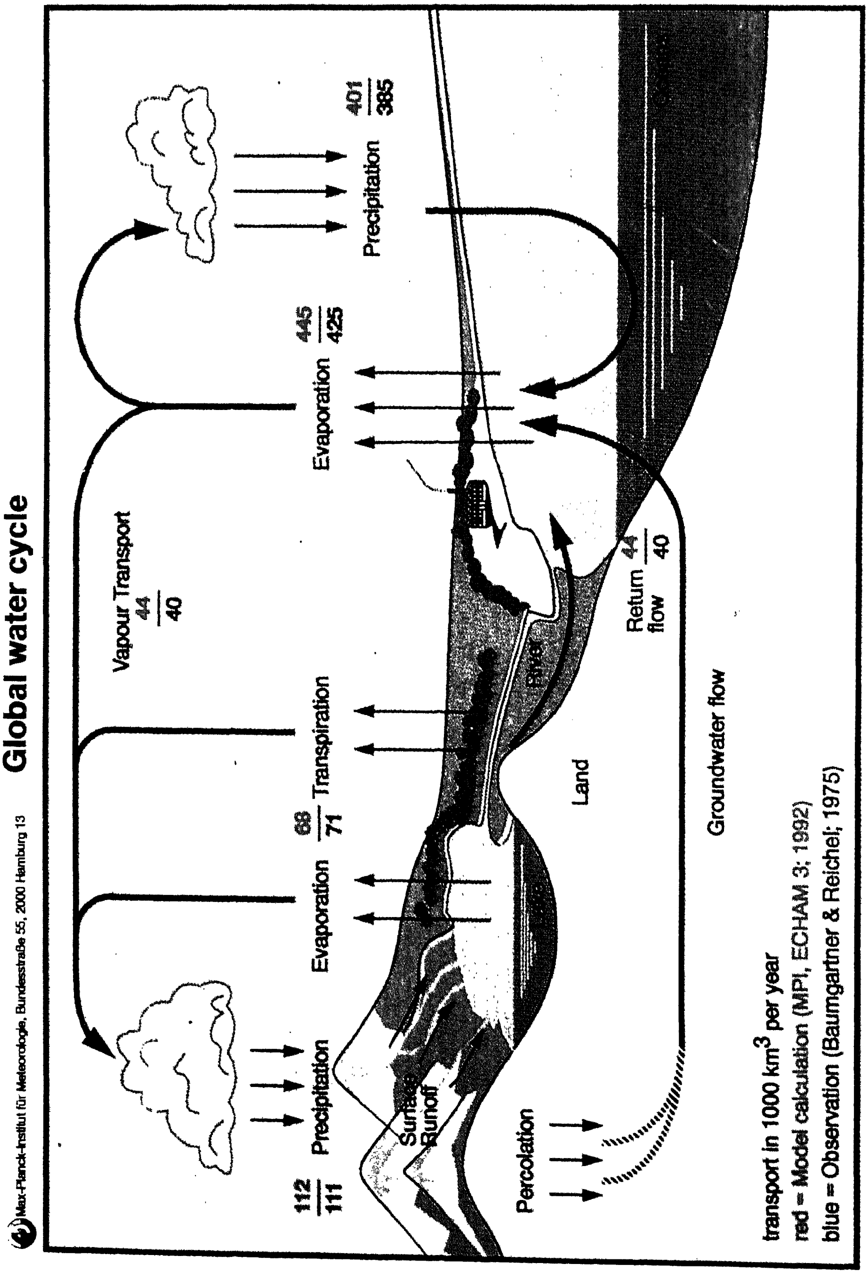


Fig 7 The annual global water cycle calculated from ECHAM3-T42 averaged over 10 years (upper figure). Evaporation and precipitation are calculated, atmospheric water transport and river return flows are obtained as residuals. Unit 1000 km^3/year.

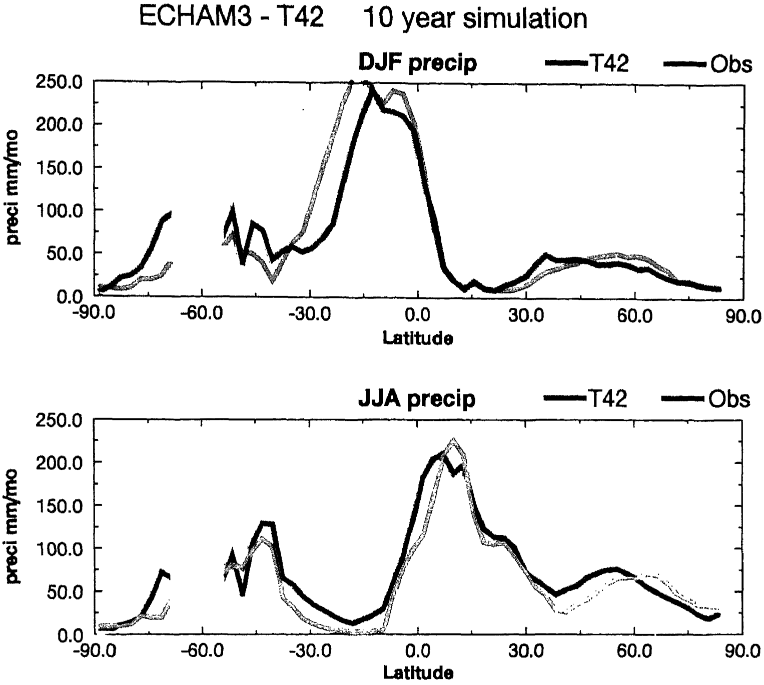


Fig 8 Precipitation over land for winter and summer respectively. Thin curve ECHAM3-T42, heavy curve climate estimate.

The agreement with the GPCC data is generally good, although the regional precipitation difficulties discussed above can also be seen from the GPCC data. There are considerable differences between the GPCC data and OLR data over land and between OLR data and SSM/I data over the oceans stressing the serious problem with the validation of precipitation.

Work is also under way to use river run-off data as an additional way to validate the hydrological cycle. A few such areas have already been investigated highlighting several interesting aspects.

For the Amazon catchment area the model is simulating the annual cycle of precipitation reasonably well, perhaps with a slightly too large amplitude. During June/July/August the precipitation amounts are very low both in reality and in the simulations but due to insufficient storage of water in the ground or in the river area in the model calculations these low precipitation rates lead to unrealistic drying of the ground with reduced evaporation and too high simulated surface temperatures.

When comparing with river runoff data from the model with observed ones for subarctic regions, for example, the MacKenzie river or the river Ob, one finds unrealistic strong runoffs during summer time. This is related to an overly high accumulation of snow especially during the spring. Where observations already show melting of snow in spring one still finds an accumulation of snow in the simulations. It is believed that the ageing of snow should lead to a reduction of the albedo. Over rough ground one would expect snow free areas which would as well lead to reductions of the albedo in snow situations. Both effects are not accounted for in the present parameterization scheme but could at least partly reduce the too late melting of snow in the area.

Land-surface Proceses

One systematic error common to several GCMs is the erroneously strong warming of the near-surface atmosphere in the Central United States in the summer. Also in the ECHAM3 model simulation too little precipitation occurs in this region, Fig. 9, and as there is plenty of energy available for evaporation, the soil dries out in the course of the integration and subsequently warms up excessively during the summer. This is not observed.

A series of experiments has begun in order to clarify whether the errors in this area are due to errors in the large-scale dynamics or if they are related to the representation of the flux of moisture from the land surface. These experiments will assess the sensitivity of the annual cycle of precipitation to the specification of the properties of the soil moisture reservoir and the general formulation of the equations governing the evaporation from bare ground or through plants.

Fig. 9 shows the results from the T21 control integration, two T21 experiments with modified boundary conditions and a T42 control integration. The atmosphere is combined with an ideal boundary condition of soil moisture (from a climatology by Mintz and Serafini, 1981)[8a] in experiment CLIS and a saturated soil (which should satisfy the atmospheric demand at all times) in experiment SATS.

In the T21 model soil wetness during the year varies between 20 and 40%, while according to climate it is 80% during the winter and spring and 20% during the summer. The model produces rainfall of the same amount as is observed in summer, but less than observed during winter and spring. Con-

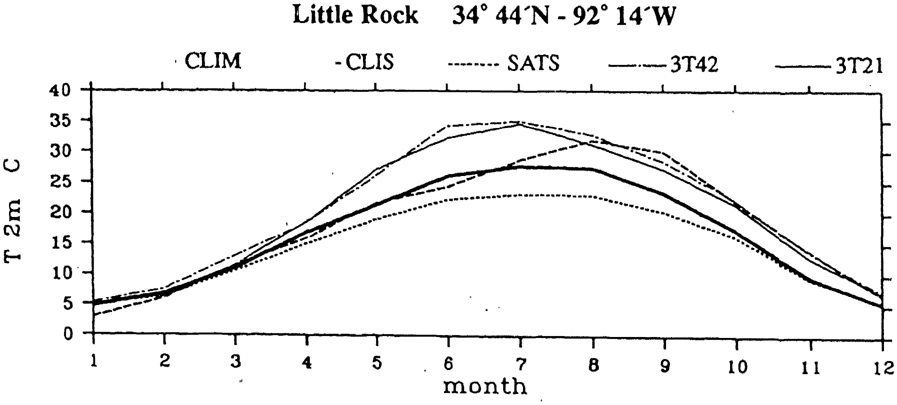


Fig 9 Annual cycle of the 2 m temperature at the model gridpoint nearest to Little Rock.
CLIM: climatology 1931-1960
3T21: from 20 years of integration with the ECHAM3 model at T21 resolution
3T42: as 3T21 but for doubled horizontal resolution.
CLIS: as 3T21 but using soil moisture boundary conditions from climatology
SATS: as 3T21 but using soil moisture boundary conditions at saturation.

sequently, the soil moisture reservoir in the model cannot be refilled which in turn is likely to suppress precipitation further. The error in the winter precipitation is both influenced by the surface boundary condition and the large-scale circulation. Precipitation increases if climate boundary conditions are applied, but is still not sufficient even if the soil is saturated. This suggests that even with correct boundary conditions the moisture transfer is not strong enough to enhance precipitation by local processes and hence, the problem is presumably related to moisture transport by the large scale circulation. This is suggested by the fact that precipitation is increased with higher horizontal resolution. Experimentation will be continued.

Low Frequency Analysis of Equatorial Zonal Wind

The dynamical properties of ECHAM3 have been verified by comparing the power spectrum of different model quantities with the corresponding analysed and observed power spectra.

Such a test was carried out for the zonal wind as observed at Singapore (Fig. 10). The "model" power spectrum was computed and averaged for two AMIP-simulations driven by the observed sea surface temperature of January 1979 to December 1988. The "analysis" power spectrum is based on ECMWF analysis data for January 1981 to December 1990. It is worthwhile to keep in mind that the analysis procedure has undergone some changes during this period. The power spectrum for these two cases was computed by a maximum entropy method using the Burg algorithm and 20 poles. These power spectra were compared with the "observation" power spectrum of the zonal wind at Singapore. The observations cover the years 1968 to 1985. The analysis is based on the same method but uses 40 poles.

Fig. 10 shows the power spectra of the zonal wind on 850mb, 200mb and 50mb, representing the dynamical behaviour near the surface, in the upper troposphere and near the top of the model. The comparison gives the following main results for the three levels:

(i) 850mb

The analyses contain only the annual cycle, while the model shows in agreement with the observed data, a low frequency peak (related to El Nino) as well. The semi-annual peak is hardly visible neither in the analysed data nor in the model simulation. (700hPa-spectra have the same characteristics).

(ii) 200mb

Analysis and model data show qualitatively the same spectrum: semi-annual and annual peak .plus the multi-annual frequency variability. Differences are presumably also due to the short data base for the analysis and the model data.

(iii) 50mb

Analyses and observations contain a distinct quasi-biannual oscillation (QBO) signal while the model is not able to reproduce this phenomenon.

Although this test is incomplete since it only includes one particular station (power spectra for the model and the analysis have been calculated for the

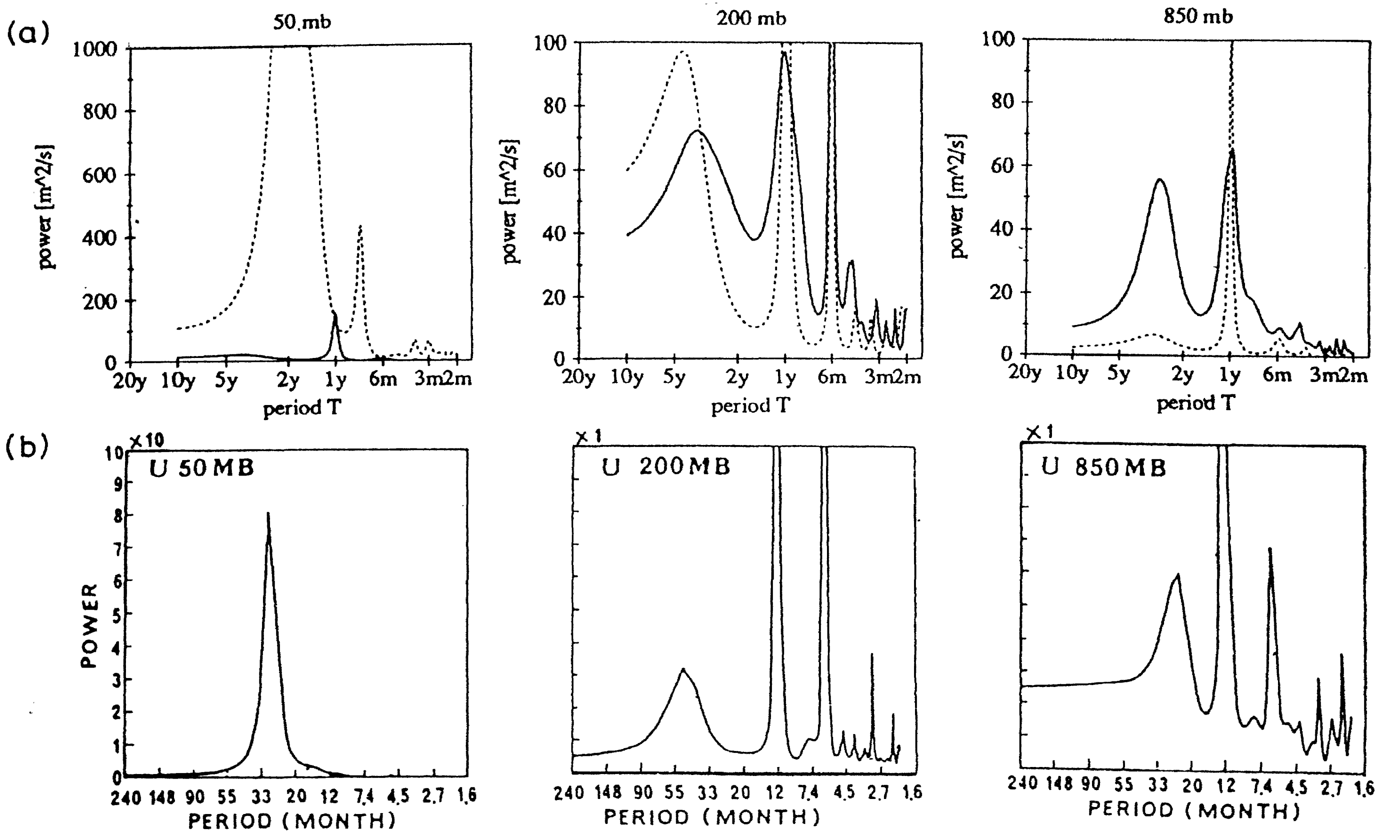


Fig 10(a) (full lines) averaged power spectrum of zonal wind of two AMIP integrations for 1979-1988. (stippled lines) power spectrum of analysed zonal wind 1981-1990.

(b) (full lines) spectrum of observed zonal wind according to Yasunari 1989 *J met Soc Japan* **67** 483-493.
]

whole equatorial domain) it shows that tropical *analysis data*, as they are available now, can show severe deviations from the dynamical behaviour of *observed data.* It is thus interesting to note that the ECHAM3 shows a multi-annual frequency variability of the tropospheric zonal wind above Singapore that agrees at least qualitatively with the observed variability. No QBO signal can be found in the model simulation.

SST-Impact Studies

The SST-anomalies during El Niño events have been found in several studies to be a major cause to anomalous circulation in the tropics. In order to make an in depth study of this forcing and to find out to what extent the ECHAM3 model could simulate it, a series of multi-year integrations using observed SST (1979-1991) and climatological SST (over the same period) have been undertaken with the following objectives:

(i) To explore the geographical distribution of the impact of SST-anomalies.
(ii) To investigate the reprodubility of this impact.
(iii) To investigate the role of El Nino SST anomalies and other tropical anomalies versus the anomalies in the North Atlantic.
(iv) What basic mechanism in the ocean/atmosphere system are responsible for the observed propagation of equatorial SST-anomalies from one ocean basin to another e.g. from the East Pacific to the Atlantic?

Preliminary results based upon an ensemble of four 10-year integrations, two using observed SST and two climatological SST, and a few additional shorter calculations for the 1982/83 event suggest the following.

There are considerable differences in the response between the tropics and extra-tropics. This can be seen in a more general way from Fig. 11 which shows the 500mb seasonal variability of JJA for the control run (using climate SSTs), the run with observed SSTs and the observed variance (from ECMWF analyses). Results from other seasons are similar. As can be seen the variance increased enforced by the observed SSTs is essentially more than doubled compared to the control run and is also in reasonably good agreement with observations.

Generally the El Nino events can be identified in all basic parameters in the tropics in particular in the upper air circulation. Precipitation anomalies are very well correlated with OLR (over oceans), and where available with *in situ* observations over land certain areas like the Sahel region and India is less good than other areas indicating more complex relation or model deficiencies. In the extra-tropics, Western Canada and United States are excellent and most major variables vary in phase with El-Niño. The signal of the European, on the other hand, is weak and here the inter-annual variability is strongly dominated by the chaotic internal variability.

The reproducibility of the result or lack of reproducibility support this result. This can be seen from Fig. 12 which shows the precipitation in two areas, central tropical Pacific and central Europe as simulated for two different runs starting from different initial conditions.

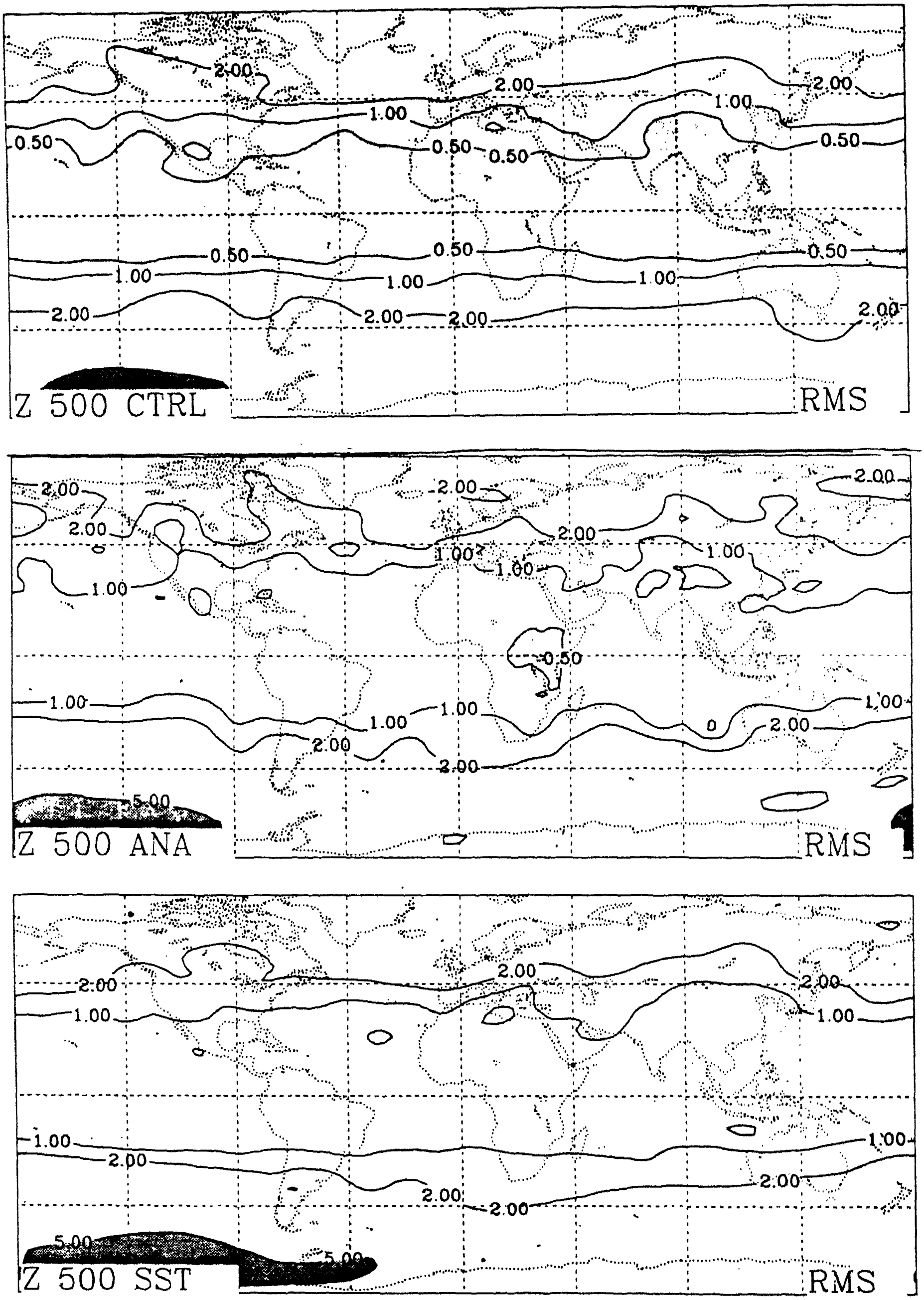


Fig 11 RMS variability in summer (JJA) at 500 mb of seasonal means 1979-1992. (a) shows the control experiment using climatological SST data (monthly averaged 1979-1992), (b) calculated from ECMWF analyses and (c) from the experiment using observed SST data.

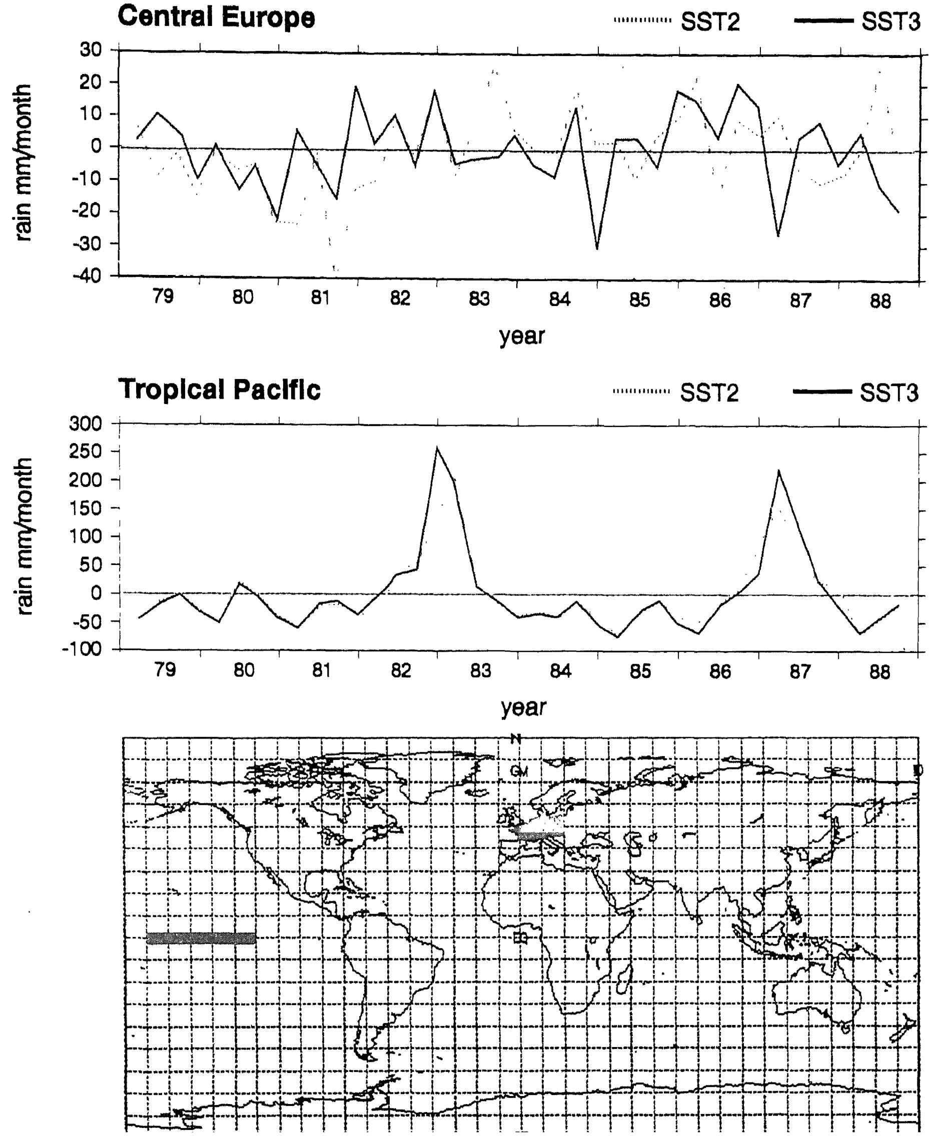


Fig 12 Simulations of seasonally-averaged precipitation anomalies over Europe and the equatorial eastern Pacific (see **lower panel**), obtained in two independent ten-year integrations with the ECHAM3 atmospheric circulation model. Observed sea-surface temperature and sea-ice conditions (as specified in the AMIP) were used in both integrations, but initial atmospheric conditions were different. Note the lack of a reliable climate signal in mid-latitudes (**upper panel**) and the high degree of reproducibility in the equatorial Pacific.

Attempts have been made to find a relation between North Atlantic SST anomalies during the period 1979-1992 and circulation anomalies. Any such relation of any meaningful amplitude has not yet been identified.

Longitude-time diagrams for tropical means of several variables show a slow eastward propagation in the experiments as well as in analysis data. The period is round 5 years. In the 200mb U field anomaly this propagation is appearing to move around the world with unchanged amplitude.while for most other variables (precipitation, OLR, vel. pot. at 850and 200mb) the features are starting over the Indian Ocean, increase in amplitude over the Pacific, with a strong reduction in amplitude over America and a fading away over the Atlantic. These features are more distinct and have larger amplitudes in the experiments but also evident in the observations. The investigation did concentrate on the period 1979 to 1992 with 3 El Niño events.

Another similar, very slow propagation but from the equator towards the poles has been found by Dickey *et al.* (1992)[9] for the angular momentum. This can also be seen in the experiments.

The investigation concentrated on the question if these apparent travelling waves are real propagations caused by an atmospheric mode or if the atmosphere passively reacts to SST anomalies which then appears like a travelling wave. Experiments with monthly mean climatological SST do not show such a propagation which suggests that the propagations in the atmosphere are only forced by the SST anomalies. Using atmospheric data from a climatological SST experiments and switching on observed SST anomalies establishes the propagating wave within a month. Also switching off the observed SST anomalies in a run during the most intense phase of El Niño lets the wave fade away within a month. So the atmosphere does not remember atmospheric anomalies for a long time and what we see therefore is in all likelihood a direct effect from the SST anomalies.

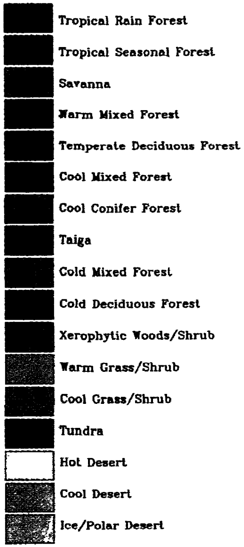
An experiment in which the atmosphere is only forced from the SST anomalies from the Pacific shows very similar propagations of a wave and similar atmospheric anomalies as the experiments with the complete SST anomalies which suggests that it is mainly the Pacific which gives a clear effect.

However, the atmospheric circulation is needed to propagate SST anomalies e.g., to allow SST anomalies to move over the American continent into the Atlantic. It can be explained by the anomalous atmospheric circulation during such events.

MONEG-Experimentation

The group has contributed to the design of the experimentation strategy of the TOGA Monsoon Experimentation group and has, been actively involved in performing and diagnosing the proposed set of numerical experiments. The aim was to intercompare the skill of 16 global atmospheric models with regard to the interannual variability of the northern hemisphere summer monsoon. Concerning the causes of this variability, the following hypothesis were tested in 90 day integrations: The model's skill in representing this variability may be influenced either by boundary conditions such as sea surface temperature anomalies or anomalies in land surface conditions or by the initial state of the atmosphere. In the beginning of the project, the example of the years 1987 (poor Indian monsoon) and 1988 (abundant Indian monsoon) was studied in order to investigate the role of the sea surface temperature anomalies in individual ocean basins. The simulation using the ECHAM3 T42 model as well as

13(a)



13 b

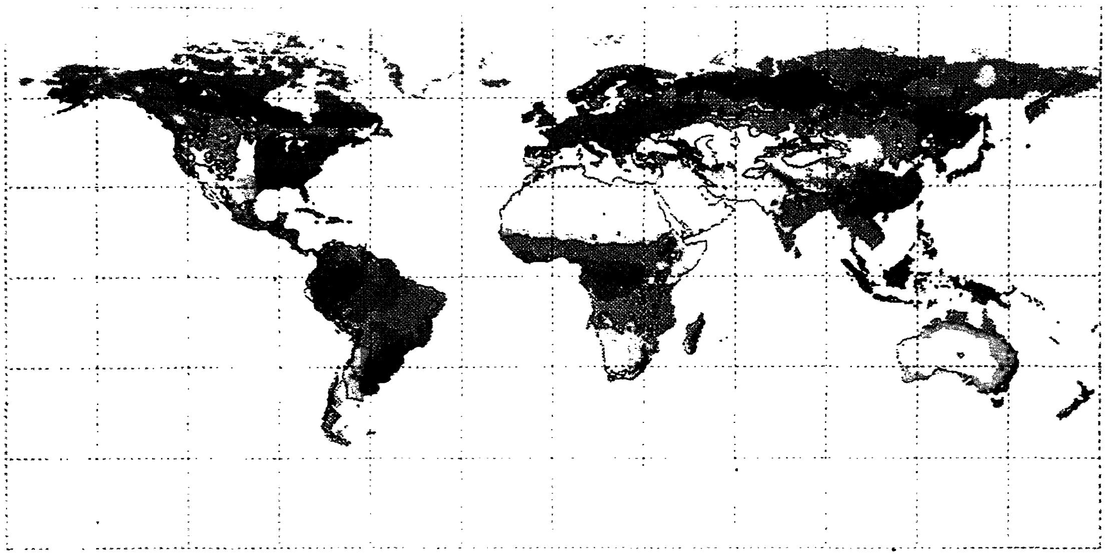
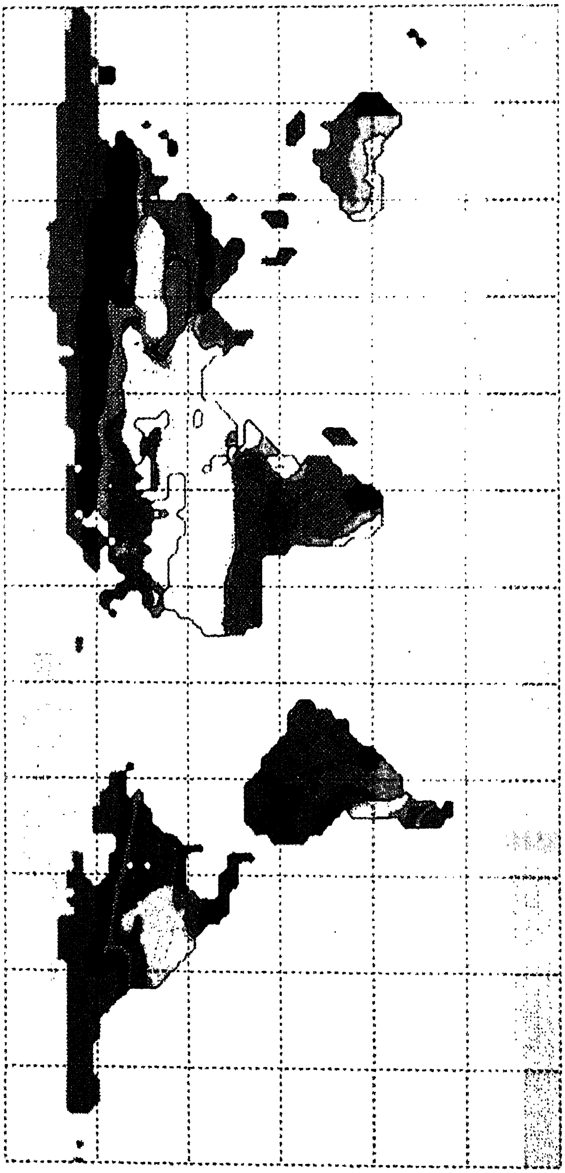


Fig 13(a) Summary of biomes used and allocation of colour to the biomes depicted in the maps.

(b) Present natural biome distribution computed from observed climatology.

those of the models showed that in this particular year the monsoon variability was forced by a relatively large Pacific Ocean sea surface temperature anomaly, with a small counteracting influence by the small anomaly in the Indian Ocean (Dümenil and Roeckner, 1992)[10]. Experimentation was continued for anomalies in the level of soil moisture over the Eurasian continent. The impact of this anomaly taken from analysed soil moisture fields from ECMWF for 1987 and 1988 is only small. Experimentation will be continued.



3(c Pre biome d d from ECHAM3-T- matolog

Biome Modelling

The *biome model* of Prentice *et al.* (1992)[11], has been used to predict global patterns of potential natural plant formations, or biomes, from climatologies simulated by ECHAM. This study was undertaken in order to show the advantage of the biome model in diagnosing the performance of a climate model and assessing effects of past and future climate changes predicted by a climate model.

Good overall agreement is found between global patterns of biomes computed from observed and simulated data of present climate (Fig. 13a-c). However there are also major discrepancies indicated by a difference in biomes in Australia, in the Kalahari Desert, and in the central part of North America. These discrepancies can be traced back to errors in simulated rain fall as well as incorrect summer or winter temperatures.

Global patterns of biomes computed from an ice age simulation reveal that North America, Europe, and Siberia should have been covered largely by tundra and taiga, whereas only small differences are seen for the tropical rain forests. A potential North-East shift of biomes is indicated from a simulation with enhanced CO_2 concentration according to the IPCC Scenario A. Little change is seen in the tropical rain forests and the Sahara. Since the biome model used is not capable of predicting changes in vegetation patterns due to a rapid climate change, the latter simulation has to be taken as a prediction of changes in conditions favourable for the existence of certain biomes, not as a prediction of a future distribution of biomes (Claussen and Esch, 1992)[12].

Climate Sensitivity Experiments

Regional Simulation

Due to the limited computational resources, projections of the future climate have been performed so far with rather coarse models, typically T21 resolution. Consequently, the confidence in the simulated regional climate change is small. An economical alternative to improve regional climate prediction is to run a high-resolution atmospheric model for a particular region or even for the whole globe but for a limited period by using the SST changes predicted by the low-resolution coupled model. This so-called time-slice technique has been applied to the last 10 years of a 100-year climate projection using the ECHAM1-T21/LSG coupled model on the basis of the IPCC scenario A (Cubasch *et al.*, 1992)[13]. The AGCM used in these time-slice experiments is a T42 version of ECHAM3. In order to separate the impact of the higher resolution (T42 *vs.* T21) from the impact of modified physics (ECHAM3 *vs.* ECHAM1), a second experiment has been performed with ECHAM3 at T21 resolution. Preliminary results are:

- The climate response of ECHAM3/T42 forced with the SST changes predicted by the ECHAM1-T21/LSG coupled model is substantiaslly different from that of the coupled model.
- The differences in the tropical precipitation signal (which is spatially more coherent in ECHAM3-T42) results mainly from a change in the convective

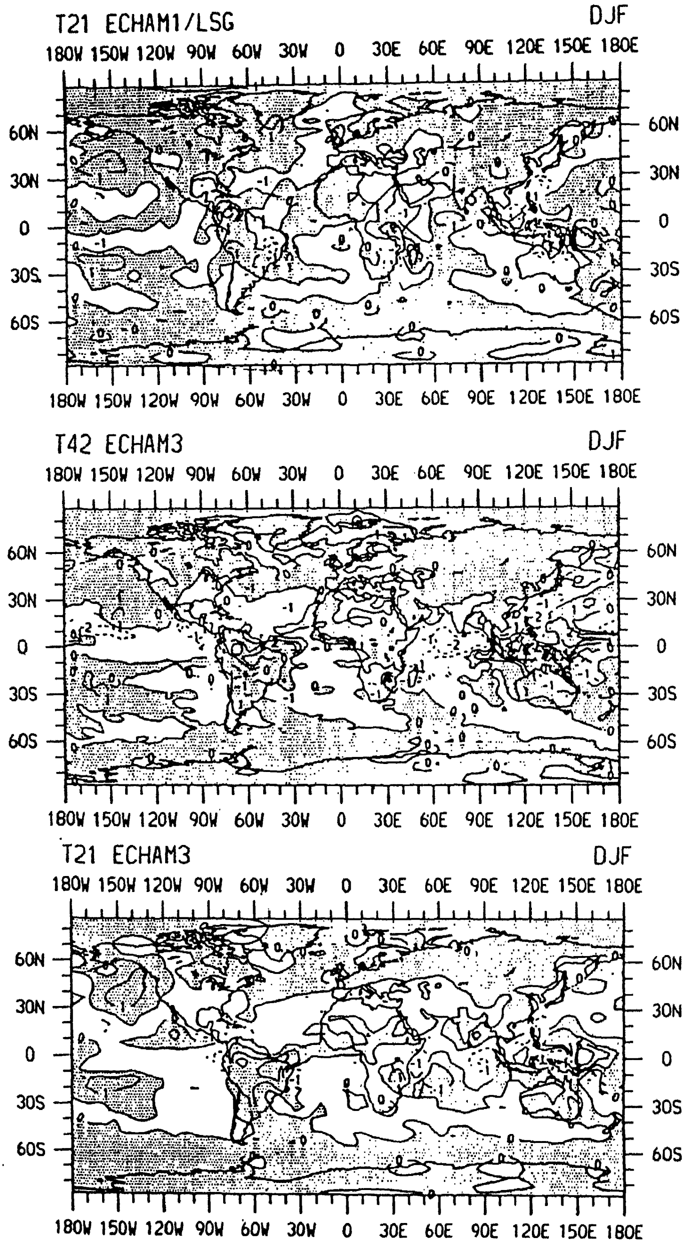


Fig 14 Change in precipitation with respect to the present winter (DJF) climate, as simulated for the last decade of a 100-year IPCC scenario A experiment.

Top: Coupled model T21|ECHAM1/LSG (10 years)

Middle and Bottom: ECHAM3 at T42 and T21 resolution, respectively forced with the SST change predicted by the couple model (30-year mean)

Contour spacing +/-1,2,4,8 mm/day. The stippling denotes areas with increased precipitation.

parameterization (mass-flux scheme in ECHAM3 *vs.* Kuo-scheme in ECHAM1) rather than from the higher resolution (Fig. 14).

- The substantial interdecadal variability in the T42 model (estimated from three 10-year realizations using the same SST forcing) does not allow to draw any firm conclusions concerning the risk of larger storm activity at middle latitudes, nor from any changes in the frequency of tropical storms.

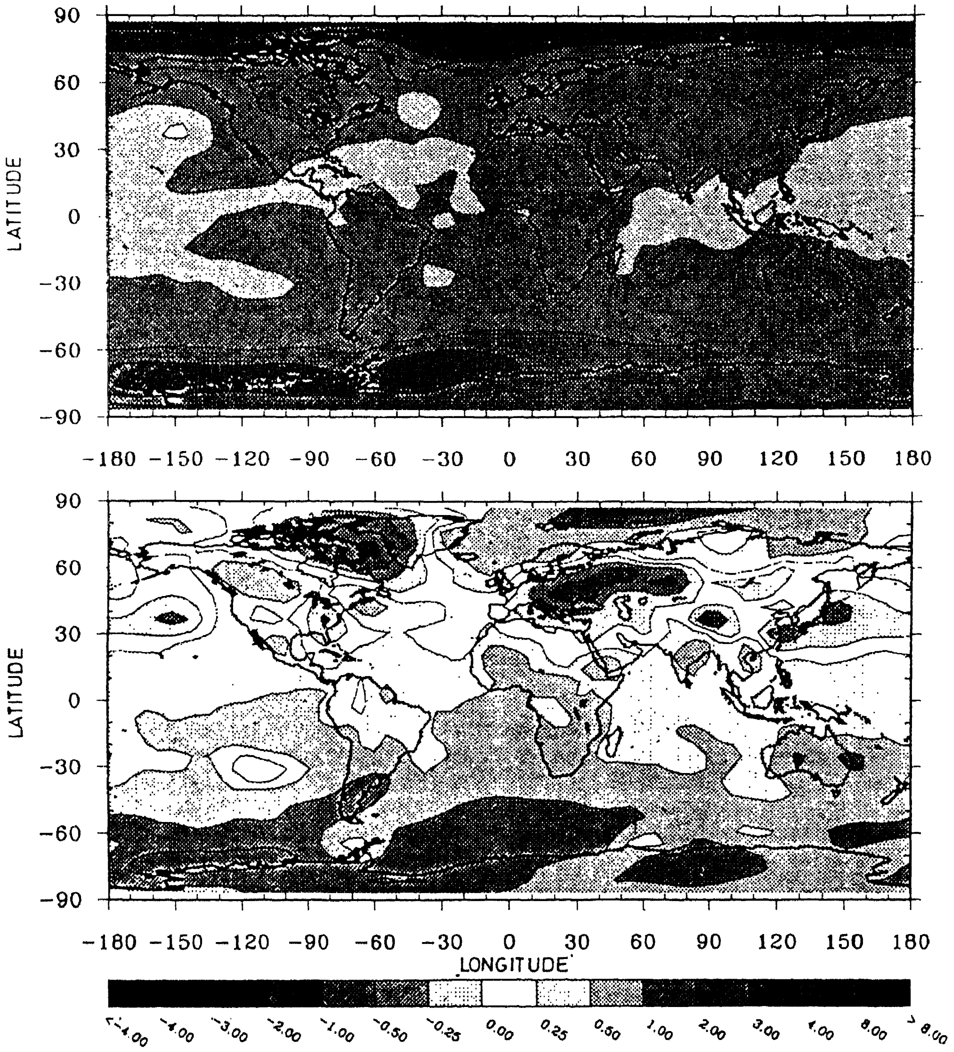


Fig 15 Equilibrium response of annual mean surface air temperature (°C) to different forcing assumptions since the beginning of the industrialization.

Top: Current greenhouse gas forcing corresponding to an effective CO^2 increase by 40% (2 Wm^{-2} in the global mean).

Bottom: Current greenhouse gas forcing as above and anthropogenic sulphate forcing (-0.7 Wm^{-2} in the global mean).

Sulphate Aerosol Forcing

In addition to atmospheric greenhouse gases, atmospheric aerosols can also affect climate. For example, biogenetic and anthropogenic sulphur species form sulphate-sulphate aerosol particles which exert a direct climatic effect due to backscattering of solar radiation and an indirect climatic effect by their ability to act as CCN (cloud condensation nuclei). Both effects tend to decrease global temperatures and may offset the global warming expected from anthropogenic greenhouse gas forcing. The hypothesis that past anthropogenic sulphur emissions have already changed our climate (Charlson *et al.*, 1991) is tested with the ECHAM1-T21 model coupled to a simple slab ocean model with a depth of 50m. In this study only the direct effect is addressed, i.e. a potential climate cooling resulting from backscattering of solar radiation by anthropogenic sulphate adopted from a simulation of the global sulphur cycle (Langner and Rodhe, 1991)[15]. The climate response is estimated by means of equilibrium runs with and without the sulphate forcing. Quasi-equilibrium is obtained after about 20 years of simulated time. In addition, a $2 \times CO_2$ forcing experiment has been performed and one which includes both effects, i.e. the current sulphate forcing and the current level of anthropogenic greenhouse gas forcing (approximately 2 W/m^2 according to IPCC estimates). The main conclusions are:

- The direct effect due to the backscattering of present-day anthropogenic sulphate aerosol is a global annual cooling of approximately 0.5°C. The cooling is stronger in the Northern Hemisphere (0.7°C) than in the Southern Hemisphere (0.3°C).
- The sulphate aerosol forcing is able to offset the current greenhouse warming over part of the Northern Hemisphere (Fig. 15).
- The equilibrium response to the sum of the current greenhouse gas forcing and anthropogenic sulphate is 0.5°C in the global annual mean, which is within the range of the observed warming during the last 100 years (0.4-0.7°C).

Global Transport Modelling

Using the semi-Lagrangian approach (Rasch and Williamson, 1990)[16], a transport code has been implemented in ECHAM3 which allows the handling of an arbitrary number of tracers. The subgrid-scale transport (vertical diffusion, convection) is treated analogously to the physical variables (heat and moisture). The code has been tested primarily with chemically inert short-lived and long-lived tracers such as radon or CFCs. Recently, the atmospheric part of the global sulphur cycle has been implemented in ECHAM. The simulation follows basically a scenario as described by Langner and Rodhe (1991)[15]: All emissions occur as DMS (dimethyl sulphide) and SO_2. The main source of SO_2 is anthropogenic emission by fossil fuel combustion and industrial activities. SO_2 is assumed to react with OH radicals producing H_2SO_4, and DMS react with OH producing SO_2. The aqueous phase oxidation of SO_2 is controlled by the oxidants H_2O_2 and O_3. SO_2 and SO_4 in clouds undergo wet removal by precipitation. The scavenging rate is assumed to be proportional to the model calculated precipitation rates. The parameterization of dry deposition of all sulphur components is based on the turbulent transfer theory used in the GCM. Since the model does not include a chemical code, the distributions of OH, H_2O_2 and O_3 have to be prescribed from a chemical model.

The results (Fig. 16) are in good agreement with those obtained from Langner and Rodhe (1991)[15]. However, a comprehensive evaluation using observational data has not yet been performed.

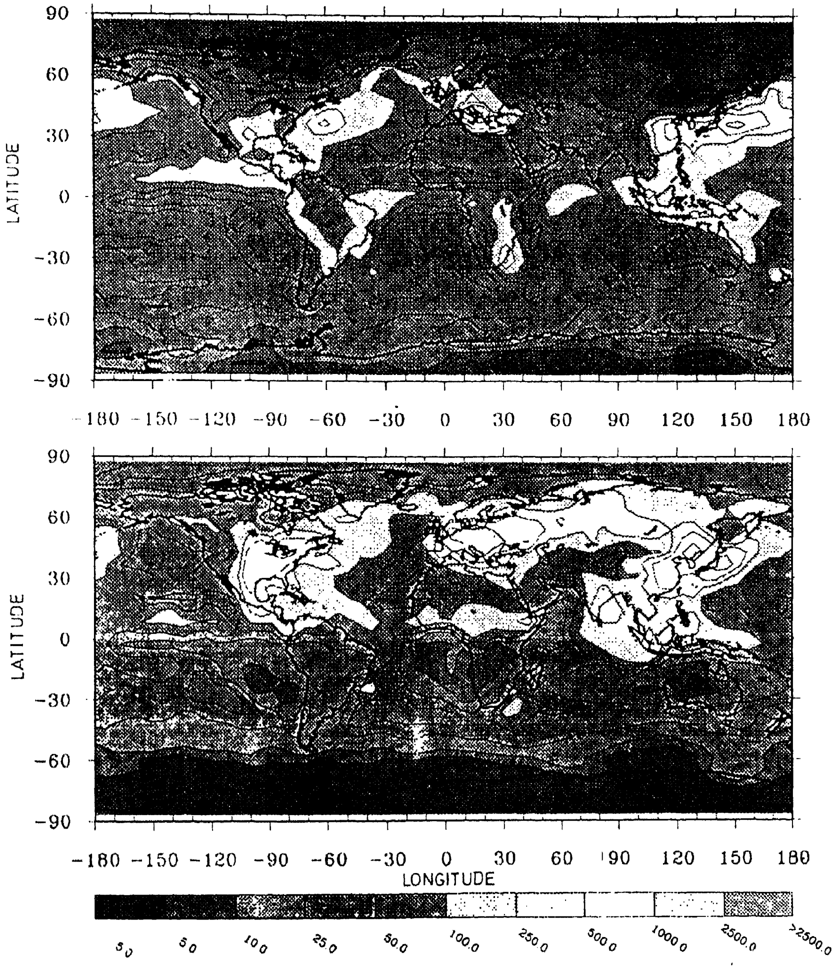


Fig 16 Dry and wet deposition of sulphur *mgS/m²*.*yr* for January (top) and July (bottom).

Acknowledgement

This article reviews the work of the staff and PhD students at MPI. We are particularly grateful for contributions by K Arpe, H Behr, R Budich, M Claussen, M Giorgetta, B Machenhauer and E Roeckner.

References

1 E Roeckner, K Arpe, L Bengtsson, S Brinkop, L Dümenil, M Esch, E Kirk, F Lunkeit, M Ponater, B Rockel, R Sausen, U Schless, S Schubert and M Windelband Simulation ofthe present-day climate with the ECHAM model: impact of model physics and resolution. *Rep* No. 93, Max-Planck-Institut für Meteorologie Hamburg (1992) 178 pp.

1 (a) E Roeckner and M Ponater In: *ECMWF Seminar Proc Tropical-Extratropical Interactions* 467-480 (1991) Reading, U.K.

1 (b) E Roeckner, M Rieland and E Keup In: *ECMWF/WCRP Workshop Clouds, Radiative Transfer and the Hydrological Cycle* (1991) 199-222, Reading, U.K.

1 (c) E Roeckner *J exp Bot* **43** (1992) 1097-1109

2 M Claussen *Scale aggregation in semi-smooth flow* *Rep* No. 87. Max-Planck-Institut für Meteorologie (1992) 21 pp. (Submitted to: *J Hydrol*)

3 M Claussen and W Klaassen *Beitr Phys Atmos* **65** (1992) 243-248

4 A Stössel and M Claussen *Rep* No. 95 Max-Planck-Institut für Meteorologie 35 pp (1992) (*Accep)ted for Climate Dyn*)

4 (a) J T Houghton, G J Jenkins and J J Ephraums *Climate Change.* The IPCC Scientific Assessment. Cambridge University Press Cambridge (1990)

5 L Dümenil and E Todini In: *Advances in Theoretical Hydrology. A Tribute to James Dooge* (*Ed.* J P O'Kane) European Geophysical Society Series on Hydrological Sciences 1 Elsevier 253 pp

6 R Sausen, S Schubert and L Dümenil *Acc. J Hydrol*

7 L Dümenil, H -J Liebscher, U Schroeder, M Schumacher and K Wilke *Rep* No. 100, Max-Planck-Institut für Meteorologie Hamburg (1993)

8 U Lohmann, R Sausen, L Bengtsson, U Cubasch, J Perlwitz and E Roeckner, 1992 *Rep* No. 92 Max-Planck-Institute für Meteorologie Hamburg (1992) 22 pp.

8 (a) Y Mintz and V Serfafini

9 J O Dickey, S L Marcus and R Hide (*Accepted by Nature*)

10 L Dümenil and E Roeckner, 1992 In: *Simulation of Interannual and Intraseasonal Monsoon Variability. Rep Workshop,* National Center for Atmospheric Research, Boulder Colorado U.S.A. (1992). World Climate Research Programme WCRP-68 WMO/TD-No. 470

11 I C Prentice, W Cramer, S P Harrison, R Leemans, R A Monserud and A M Solomon *J Biogeograph* **19** (1992) 117-134

12 M Claussen and M Esch *Rep* No. 89, Max-Planck-Institut für Meteorologie (1992) 19 pp. (*Acc publ Climate Dyn*)

13 U Cubasch, K Hasselmann, H Höck, E Maier-Reimer, U Milolajewicz, B D Santer and R Sausen *Climate Dyn* **8** (1992) 55-69

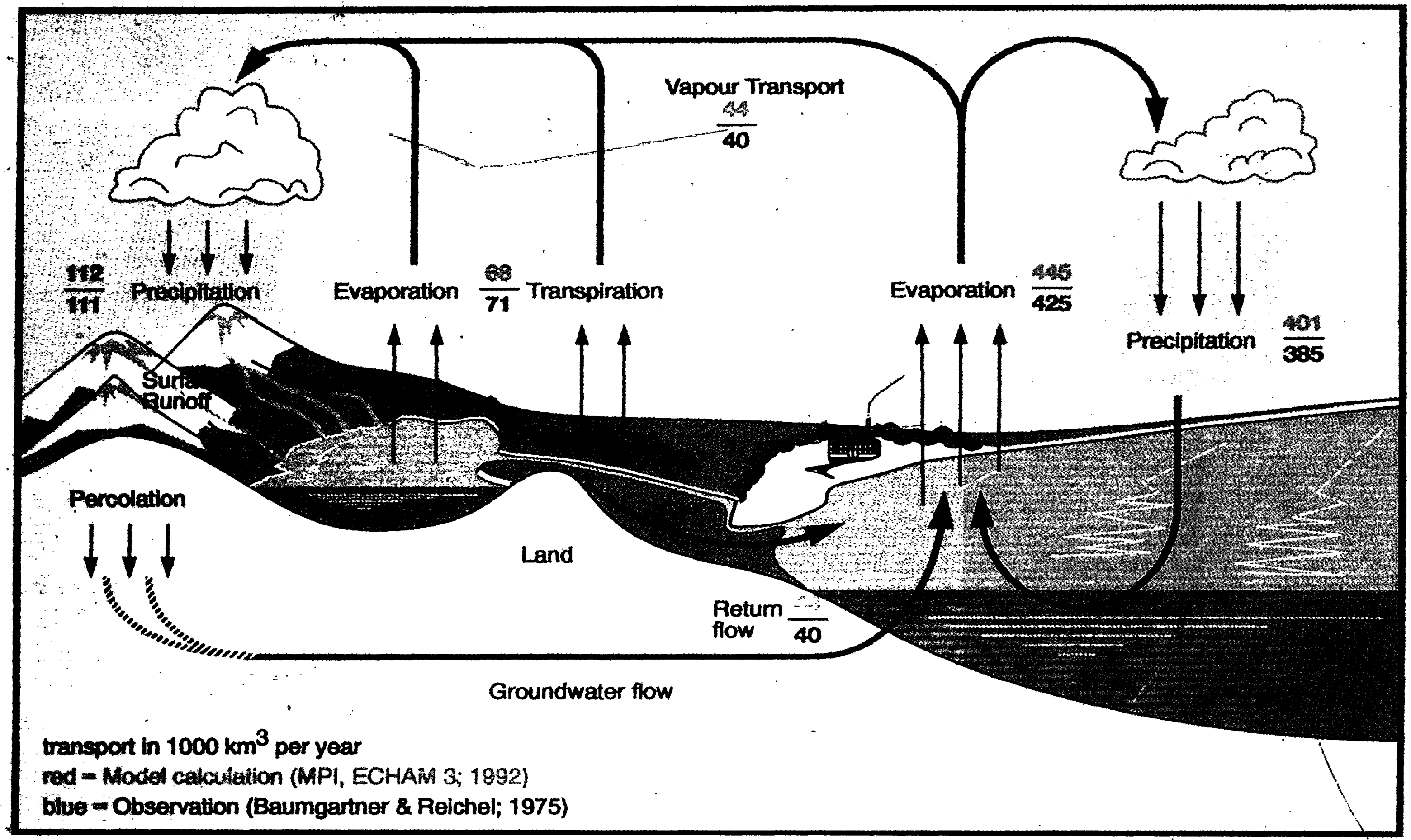
14 R J Charlson, J Langner, H Rodhe, C B Leovy and S G Warren *Tellus* **3** AB (1991) 152-163

15 J Langner and H Rodhe *J atmos Chem* **13** (1991) 225-263

16 P J Rasch and D Williamson *Models of the atmosphere, QJR Meteorol Soc* **116** (1990) 1071-1090

Max-Planck-Institut für Meteorologie, Bundesstraße 55, 2000 Hamburg 13

Global water cycle



transport in 1000 km^3 per year

red = Model calculation (MPI, ECHAM 3; 1992)

blue = Observation (Baumgartner & Reichel; 1975)